PROBABILISTIC ANALYSIS OF THE COMPRESSIBILITY OF SOILS

A Dissertation

by

BYOUNG CHAN JUNG

Submitted to the Office of Graduate Studies of
Texas A&M University
in partial fulfillment of the requirements for the degree of

DOCTOR OF PHILOSOPHY

May 2009

Major Subject: Civil Engineering

PROBABILISTIC ANALYSIS OF THE COMPRESSIBILITY OF SOILS

A Dissertation

by

BYOUNG CHAN JUNG

Submitted to the Office of Graduate Studies of
Texas A&M University
in partial fulfillment of the requirements for the degree of

DOCTOR OF PHILOSOPHY

Approved by:

Co-Chairs of Committee,	Giovanna Biscontin
	Paolo Gardoni
Committee Members,	Charles Aubeny
	Jerome J. Schubert
Head of Department,	David V. Rosowsky

May 2009

Major Subject: Civil Engineering

ABSTRACT

Probabilistic Analysis of the Compressibility of Soils. (May 2009)

Byoung Chan Jung, B.S., Dae Jeon University;

M.S., Texas A&M University

Co-Chairs of Advisory Committee: Dr. Giovanna Biscontin
Dr. Paolo Gardoni

Geotechnical engineers are always faced with uncertainties and spatial variations in material parameters. In this work, we propose to develop a framework able to account for different types of uncertainties in a formal and logical manner, to incorporate all available sources of information, and to integrate the uncertainty in an estimate of the probability.

In geotechnical engineering, current soil classification charts based on CPT data may not provide an accurate prediction of soil type, even though soil classification is an essential component in the design process. As a cheaper and faster alternative to sample retrieval and testing, field methods such as the cone penetration test (CPT) can be used. A probabilistic soil classification approach is proposed here to improve soil classification based on CPT. The proposed approach provides a simple and straightforward tool that allows updating the soil classification charts based on site-specific data.

In general, settlements can be the result of surface loads or variable soil deposits. In current practice, the analysis to determine settlements is deterministic. It assumes that

the soil profile at a site is uniform from location to location, and only allows limited consideration of the variations of the material properties and initial conditions within soil layers in spite of the wide range of compositions, gradations, and water contents in natural soils. A Bayesian methodology is used to develop an unbiased probabilistic model that accurately predicts the settlements and accounts for all the prevailing uncertainties. The proposed probabilistic model is used to estimate the settlements of the foundation of a structure in the Venice Lagoon, Italy. The conditional probability (fragility) of exceeding a specified settlement threshold for a given vertical pressure is estimated. A predictive fragility and confidence intervals are developed with special attention given to the treatment and quantification of aleatory and epistemic uncertainties. Sensitivity and importance measures are computed to identify the key parameters and random variables in the model.

ACKNOWLEDGMENTS

First, I owe to my sincere gratitude to God.

I would like to express my deepest gratitude to my advisors, Dr. Giovanna Biscontin, and Dr. Paolo Gardoni, who have advised and supported me with encouragement, valuable suggestions, challenges, and patience on this work. I would like to also thank Dr. Charles Aubeny and Dr. Jerome J. Schubert for serving as the advisory committee.

In addition, I greatly acknowledge the support and endless encouragement of my parents, who always showed open hearts and supported me unconditionally over the past several years of this work. Especially, this dissertation is dedicated to my parents, Sun Ho Kim, Soon Tae Jung, and Hae Bok Lee.

Finally, thanks to my wife, Jung Mi Do, and sweet daughter, Ellie Esther Jung, for their trust, understanding, and patience during this demanding part of my life.

TABLE OF CONTENTS

CHAPTER Page

CHAPTER Page

LIST OF FIGURES

FIGURE Page

FIGURE Page

LIST OF TABLES

TABLE Page

CHAPTER I

INTRODUCTION

1.1. OVERVIEW

In recent years, a reliability analysis in geotechnical projects has increased. Many applications are developed to provide theoretical and explicit quantification for practical designs and construction problems. The common geotechnical projects are different in many ways form the other fields. However, the probabilistic methods in geotechnical engineering have a little applications and methodologies. The geotechnical engineers must deal with uncertainties in material parameters and spatial variations. Therefore, we propose to develop a framework able to account for different types of uncertainties in a formal and logical manner, to incorporate all available sources of information, and to integrate the uncertainty in an estimate of the probability.

In geotechnical engineering, soil identification is an essential component in the design process. As a cheaper and faster alternative to sample retrieval and testing, field methods such as the cone penetration test (CPT) can be used. Unfortunately, current soil classification charts based on CPT data may not provide an accurate prediction of soil type. A probabilistic soil classification approach is proposed here to improve soil classification based on CPT. The proposed approach provides a simple and

---

This dissertation follows the style and format of the *Journal of Geotechnical and Geoenviromental Engineering*.

straightforward tool that allows updating the soil classification charts based on site-specific data. The updated classification chart can be used for a more accurate classification of the soil and accounts both for prior information available before conducting the tests and for the site-specific data.

Excess settlements are a serious problem for many structures and can cause significant damage. In general, settlements can be the result of surface loads or variable soil deposits. Because there are no accurate and reliable models to estimate differential settlements, mitigation methods have been developed to minimize the problem, such as fill pads, pre-loading, or the use of deep foundations. For some structures, such as the Kansai airport in Japan, the solution is still more complex and inventive, deferential settlements are counteracted by raising or shortening particular columns supporting the building. When the preventive measures against excess settlement are inexistent or insufficient, repairs after construction can be extremely expensive or lead to the demise of the structure (Poulos, 2005).

Uncertainties are introduced in the reliability frameworks through the natural variability of soils, the limited number of borings and tests usually conducted on soil samples, the scatter in the estimated soil properties, and the limitations of the methods. The probabilistic tools provide a sound basis for decisions about the need to mitigate for methodology and the type of intervention most suitable for the purpose.

Any model needs to be assessed and validated against data. This research propose to use the existing and extensive database on Venice Lagoon soil properties and the field data obtained from the full-scale test embankment constructed near Venice,

Italy. The goal of this project is to measure directly in situ the stress-strain-time properties of Venetian soils. The importance of preserving the historic city of Venice from increasing flooding has resulted in the currently favored plan to construct tilting flap gates at the three inlets of the lagoon (Harleman et al., 2000). The gates' foundations are extremely sensitive to settlements and the Italian project aims at developing deeper understanding of the mechanisms leading to settlements and ultimately better predictions of differential settlements. This task is particularly difficult because of the natural variability in Venice soil deposits, characterized by similar mineralogy, but extremely erratic depositional patterns.

In this context, the development of a framework such as that proposed here is an invaluable tool in identifying the major sources of uncertainty, in quantifying the accuracy. The proposed project will have access to a high-profile and an extremely well documented case study, perfectly suited for the validation of the proposed models.

1.2. OBJECTIVES

This research focuses on the reliability analysis applied to Venice soil. In particular, the following objectives will be addressed:

- ***Objective 1: Develop a probabilistic framework for soil identification***

  Identify the soil class with in-situ data using Cone Penetration Tests (CPT) to improve the accuracy of classification.

- ***Objective 2: Develop a probabilistic framework to predict settlements***

Develop a framework for handling the uncertainty in geotechnical data and selected geotechnical models for settlement prediction. Probabilistic models to provide an accurate and unbiased model will be developed to account for the uncertainties.

- ***Objective 3: Assess fragility estimates for settlement***

  Develop a reliability framework to assess the probability that a specified threshold settlement is exceeded.

- ***Objective 4: Assess fragility estimates for time-dependent settlement***

  Estimate the time-dependent settlement and develop a reliability framework to assess the fragility estimates.

1.3. ORGANIZATION OF DISSERTATION

The dissertation is composed of the five chapters, each containing a journal paper.

In Chapter II, probabilistic soil identification is developed. The soil identification is accomplished by tailoring the CPT classification chart to a specific site, or region, through the development of a local database of CPT tests and associated laboratory classifications. The title of the corresponding paper is "Probabilistic soil identification based on Cone Penetration Tests" and was published in the *Geotechnique*, 58(7).

In Chapter III, probabilistic soil compression model is developed using a probabilistic method that properly accounts for the prevailing uncertainties, including model errors and statistical uncertainty. The probabilistic model is based on a unified

soil compression model developed by Biscontin et al. (2007). The corresponding paper titled "Bayesian updating of a unified soil compression model" will be published in the journal *Gerisk* for a special publication.

In Chapter IV, a new probabilistic soil compression model is proposed to predict settlement of foundations. The proposed model is updated from an available probabilistic model in Chapter III. The developed probabilistic model is used to assess the settlement and fragility estimates. Also sensitivity and importance measures are conducted to investigate the contribution of variables. The corresponding paper titled "Reliability analysis of settlements" is currently under preparation for submission.

In Chapter V, Reliability analysis for time-dependent settlement is conducted using a probabilistic compression model in Chapter IV. The corresponding paper titled "Reliability analysis of time-dependent settlements" is currently under preparation for submission.

Finally, in Chapter VI, the conclusions are included.

CHAPTER II

PROBABILISTIC SOIL IDENTIFICATION BASED ON CONE PENETRATION TESTS

2.1. INTRODUCTION

In geotechnical engineering, soil classification and identification of geological units are extremely important steps in the development of any foundation, embankment or excavation project. Typically, samples are retrieved from a site and tested in the laboratory to determine grain size distribution and consistency limits, as used for classification purposes. Among in situ testing methods, the cone penetration test is widely accepted to identify soil type and define subsurface profile. The test is economical and provides continuous records with depth; however it does not allow sample retrieval for observation and laboratory testing.

Soil classification by CPT is based on empirical charts developed by, among others, Douglas and Olsen (1981), Robertson et al. (1986), and Robertson (1990). While the classification charts are widely applicable, in some cases they are generic and may fail to provide accurate soil classification.

Since faulty and misleading classifications can result in costly additions or modifications to a project, it is important to improve the accuracy of CPT classification. This can be accomplished by tailoring the CPT classification chart to a specific site, or region, through the development of a local database of CPT tests and associated laboratory classifications. The site-specific chart would not be limited to a single project

and therefore somewhat limited in its usefulness- but could be used for the geographical area in which the soil units of interest are present.

A method must be devised to update the existing charts to account for the information provided in the site-specific database. New additional data should also be incorporated as the information becomes available. Furthermore, since many geotechnical offices tend to work in the same region, geotechnical engineers develop a general knowledge of the soils in their area, often described as "engineering judgment". The method should also be able to incorporate engineering judgment in the development of the regional identification chart. Finally, given the inherent uncertainties in the classification estimates (unless a sample is retrieved and tested), a measure of the accuracy of the classification will also be a valuable contribution to CPT use.

Zhang & Tumay (1996) explored the accuracy of CPT classification through the use of conformal mapping of two independent indices: soil classification index and soil in situ state index. A fuzzy subset approach was later introduced (Zhang & Tumay, 1999). In these methods based on statistical approaches, the uncertainties are attributed to both the soil mechanical behaviour and soil composition.

The proposed approach addresses the concerns listed above by developing a Bayesian updating framework to develop site-specific CPT classification charts. In addition, it also properly accounts for the prevailing uncertainties and can be used to assess the accuracy of the soil classification. The first step is to identify relevant soil classes and simplify the selected CPT classification chart in order to develop a rational updating method based on site-specific information and tests. The second step involves

the development of a Bayesian formulation to tailor the soil identification to site-specific CPT data.

The following section provides a brief overview of classification charts with CPT. The proposed method for developing site-specific soil identification charts is subsequently described. Finally, the proposed methodology is illustrated developing two site-specific charts using data from two sites, the Treporti Test Site (TTS) near Venice, Italy and the National Geotechnical Experimentation Sites (NGES) in College Station, Texas. The applicability of the first site-specific chart for other sites in Venice Lagoon is assessed using data from the Malamocco test site, approximately 20 km from TTS.

2.2. ANALYTICAL FORMULATION OF SOIL CLASSIFICATION CHARTS

Among the many different CPT soil classification charts, Robertson's chart (1990) is one of the most widely used in geotechnical applications. This chart has two axes: the normalized friction ratio, $F_R = 100 f_s / (q_t - \sigma_{v0})$, and the normalized tip resistance, $Q_t = (q_t - \sigma_{v0}) / \sigma'_{v0}$, from CPT testing. A soil is classified based on the particular combination of F_R and Q_t measured during in situ testing. Robertson's chart is divided into nine areas corresponding to different soil types as shown Fig. 2.1. Table 2.1 lists the description of each soil type.

A simple mathematical description of the boundaries between soil classes in Robertson's classification chart is needed to develop a rational method for modifying and updating the chart based on site-specific data. Since these boundaries are not defined by analytical relationships, exponential curves are fitted to the existing set of

Table 2.1. Description of soil classes from Robertson's soil classification chart (adapted from Robertson, 1990)

Area	Soil description
1	Sensitive, fine-grained soils
2	Organic soils and peat
3	Clays [clay to silty clay]
4	Silt mixtures [silty clay to clayey silt]
5	Sand mixtures [sandy silt to silty sand]
6	Sand [silty sand to clean sand]
7	Sand to gravelly sand
8	Sandy-clayey sand to "very stiff" sand
9	Very stiff, fine-grained, overconsolidated or cemented soil

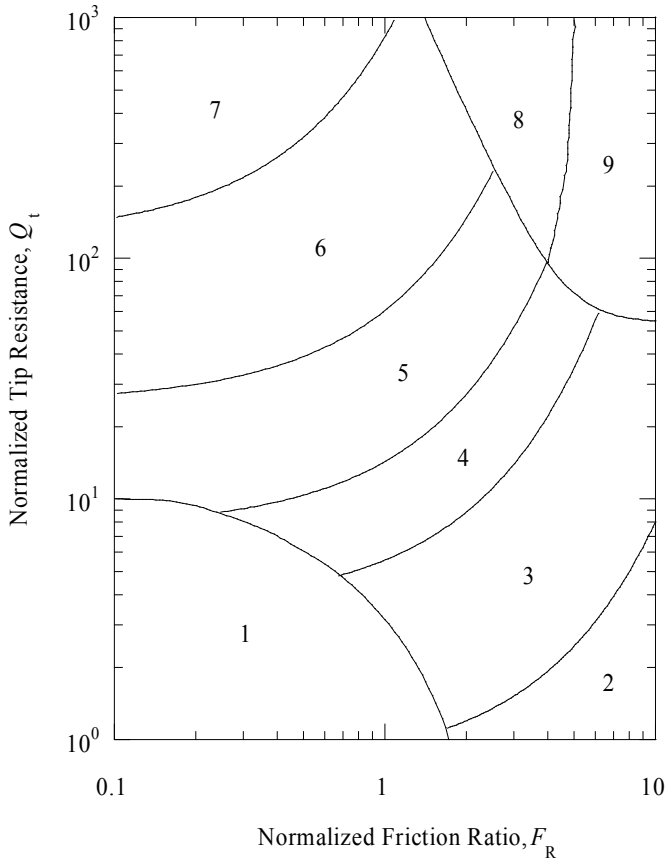


Fig. 2.1. Soil classification chart (adapted from Robertson, 1990)

boundaries. Each boundary can then be easily transformed into a linear relationship by simply transferring to a semi-logarithmic space.

$$Q_t = b_i e^{m_i F_R} \tag{2.1}$$

$$\ln(Q_t) = \ln(b_i) + m_i F_R = a_i + m_i F_R \tag{2.2}$$

In the updating process described next, the basic shape of the boundaries between classes is assumed to remain the same as in the original chart, while allowing the actual position of the boundaries to vary.

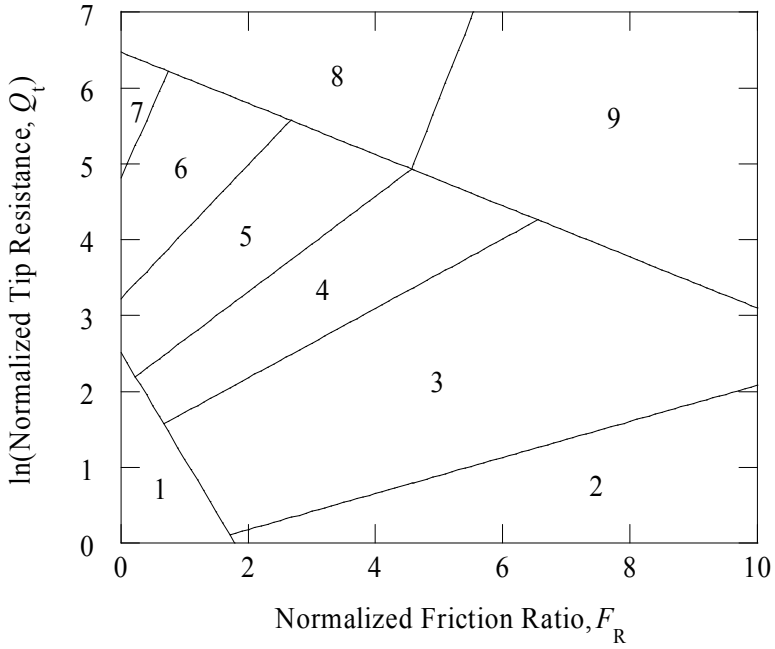


Fig. 2.2. Soil classification transformed chart

Fig. 2.2 shows the chart in the semi-logarithmic space after the transformation from Robertson's original chart with Eqs. (2.1) and (2.2). The values of a_i and m_i are

determined to fit the empirical chart developed by Robertson as listed in Table 2.2. In the next section, the procedure for updating a_i and m_i based on site-specific data is discussed.

Table 2.2. Values of boundary variables from Robertson's chart

Boundary	a_i	m_i
1	2.513	–1.402
2	–0.296	0.238
3	1.265	0.457
4	2.053	0.630
5	3.224	0.881
6	4.808	1.907
7	6.472	–0.337
8		2.130

2.3. PROBABILISTIC SOIL CLASSIFICATION

This section focuses on the development of a Bayesian probabilistic formulation to update the classification chart with site-specific field and laboratory data. Soil samples and CPT records from relatively close locations are needed in order to ensure the information refers to the same soil. The samples are tested in the laboratory to establish the classification (or, alternatively, class) of the soil. Laboratory classification and CPT tip resistance and friction ratio at the same depth are referred to as field/laboratory data pairs in the following sections. The laboratory data might be inexact due to errors in the measurement devices or procedures; however the statistics of the measurement errors can typically be obtained from calibration and/or engineering judgment (Gardoni et al.,

2002b). In this section, the approach that considers exact measurements is developed first. Next, measurement errors are introduced in the formulation.

2.3.1. Soil Classification with Exact Measurements

To implement a probabilistic soil classification, the following conditional probability is defined

$$\pi_{\alpha|\beta} = p\left(\text{soil class} = \alpha \,|\, \text{area} = \beta\right) \tag{2.3}$$

as the probability that the soil belongs to class α, given that F_R and Q_t are in area β of the classification chart. A class is essentially the description of a soil type. This can be, for example, any of Robertson's soil types, one of the groups recognised by the Unified Soil Classification System (USCS), or an ad-hoc definition, whichever is most appropriate for the site or project. Separate classes may be defined in the event of two soils with different responses, but similar traditional classification.

Since the soil classes are mutually exclusive and collectively exhaustive, the conditional probabilities for all soil classes sum up to 1, that is

$$\sum_{\alpha} \pi_{\alpha|\beta} = 1 \tag{2.4}$$

The overall goals of the probabilistic formulation are to

- estimate the probability of having any of the soil classes, given F_R and Q_t are in a certain area; and
- estimate the most likely location of the boundaries between each area.

Considering Robertson's chart in Fig. 2.1, the boundary between areas 5-8 and 4-9 is regarded as continuous, since it marks the transition between prevalently fine-grained soils and mostly sandy soils. However, the two segments are defined independently (i.e., the linear segments in Fig. 2.2 can have different slopes). Therefore, 15 unknown boundary variables (one intercept and one slope for each of the boundary lines, except for the boundary between areas 8 and 9 that only has an unknown slope) and 81 conditional probabilities, corresponding to the nine soil classes and the nine areas, need to be estimated.

It is foreseeable that a particular site or even region will only have a limited number of soil classes present. Therefore, on a case-by-case basis the actual number of variables can be smaller, depending on the soil classes involved. It may also be convenient to group certain classes and form a new soil class encompassing more than one of the original classes. As more data become available, more classes can be defined to better differentiate between soils as desired.

If, as an example, only 3 classes of soils are present (i.e., clay, silt mixtures, and sand mixtures in Robertson's chart) as shown in Fig. 2.3, the model has only four unknown boundary variables (one intercept and one slope for each of the boundary lines) and nine conditional probabilities (listed in Table 2.3) that need to be estimated.

Fig. 2.4 shows all thirteen unknown parameters. However, not all values of boundary variables a_2 and m_2 are acceptable, since the boundary line they define is not allowed to cross any of the other boundary lines with a positive slope. Therefore, the zero intercept of the boundary line between areas A and B must be greater than or equal

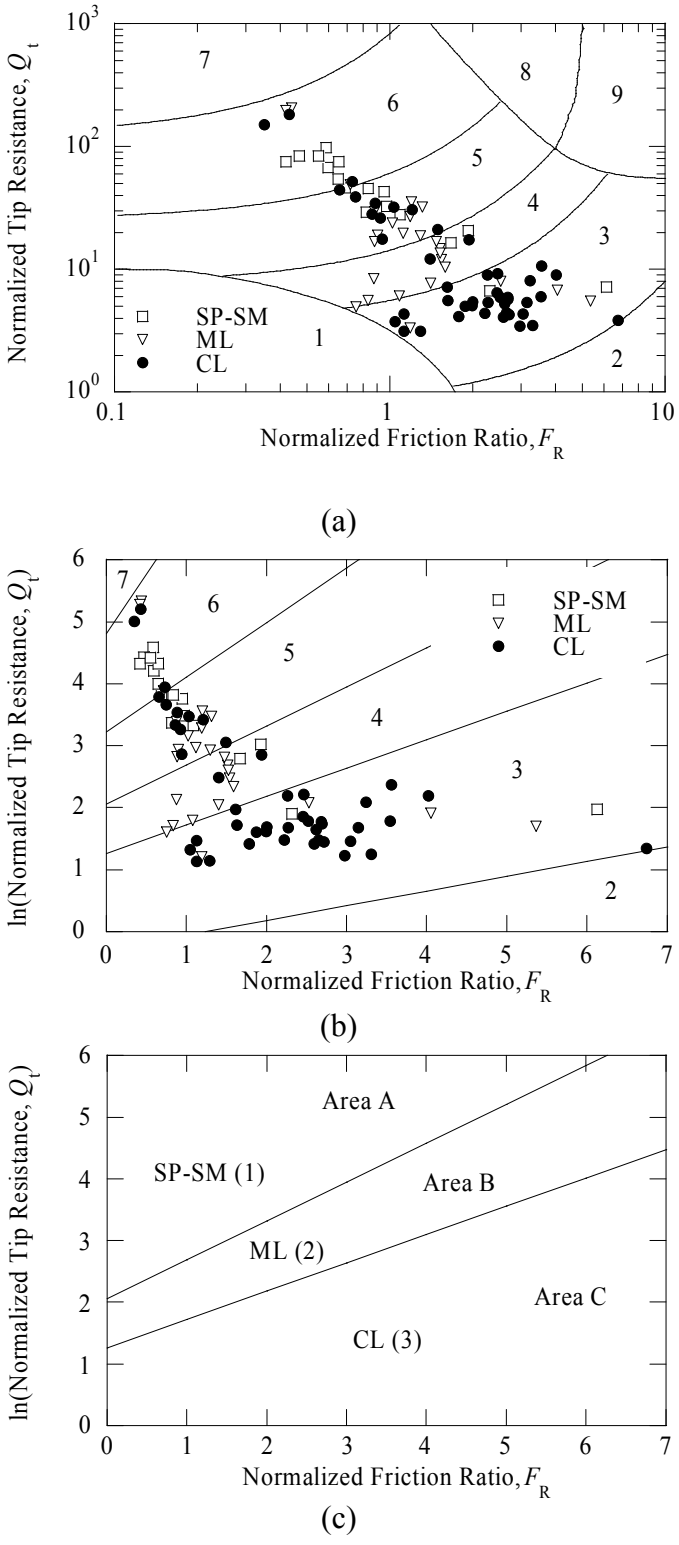


Fig. 2.3. Transformed Robertson's classification chart with soil data and simplified chart

Table 2.3. Conditional probabilities based on soil classes and areas

Soil classes	Areas		
	A	B	C
SP-SM (1)	$\pi_{1\|A}$	$\pi_{1\|B}$	$\pi_{1\|C}$
ML (2)	$\pi_{2\|A}$	$\pi_{2\|B}$	$\pi_{2\|C}$
CL (3)	$\pi_{3\|A}$	$\pi_{3\|B}$	$\pi_{3\|B}$

to the intercept of the boundary line between areas B and C, that is $a_2 \geq a_1$. Similarly, the slopes of all boundary lines in this example must be positive and such that $m_2 \geq m_1$. These constraints are imposed in the proposed approach by defining a_2 and m_2 as

$$a_2 = a_1 + \theta_a, \quad m_2 = m_1 + \theta_m \tag{2.5}$$

where θ_a and θ_m are non-negative quantities.

In this example, Eq. (2.4) can be rewritten as

$$\pi_{3|j} = 1 - \left[\pi_{1|j} + \pi_{2|j} \right] \tag{2.6}$$

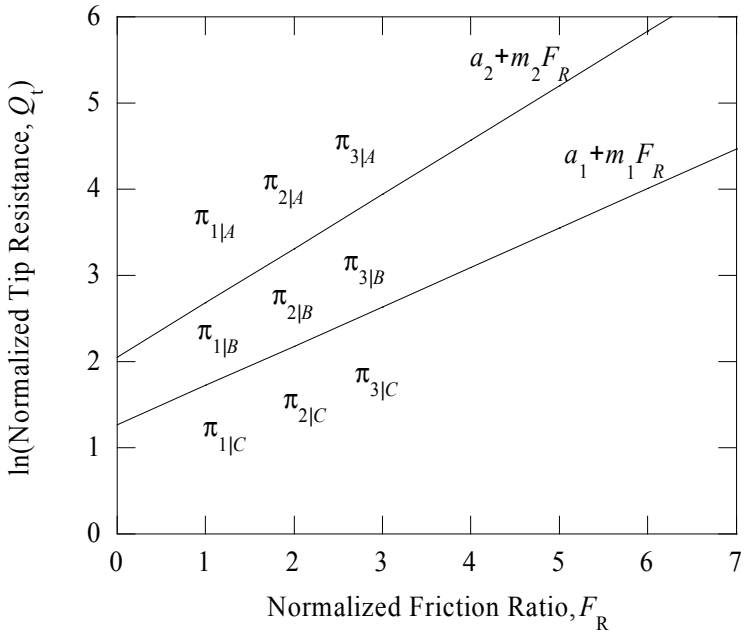


Fig. 2.4. Sets of model parameters in identification chart

2.3.2. Soil Classification with Imprecise Measurements

The proposed probabilistic approach for soil identification can also account for imprecise measurements, when a soil might be misclassified based on incorrect results from laboratory tests. To account for possible misclassification, the following quantity is defined

$$\pi_{\mathrm{I}\alpha|\gamma,\beta} = p\left(\text{lab test} = \mathrm{I}\alpha \,|\, \text{soil class} = \gamma, \text{area} = \beta\right) \tag{2.7}$$

as the probability that the laboratory test indicates that the soil belongs to class α, given that the soil in reality belongs to class γ, and F_R and Q_t are in area β.

The total probability rule (Ang and Tang, 2006) can then be used to compute the conditional probability

$$\pi_{\mathrm{I}\alpha|\beta} = \sum_{\gamma}\left[\pi_{\mathrm{I}\alpha|(\gamma,\beta)} \cdot \pi_{\gamma|\beta}\right] \tag{2.8}$$

where $\pi_{\mathrm{I}\alpha|\beta}$ represents the probability that the laboratory test indicates soil class α, given that F_R and Q_t are in area β.

As a result, the independent unknown model parameters are only ten $\boldsymbol{\Theta} = (a_1, m_1, \theta_a, \theta_m, \pi_{1|A}, \pi_{1|B}, \pi_{1|C}, \pi_{2|A}, \pi_{2|B}, \pi_{2|C})$. The independent unknown model parameters $\boldsymbol{\Theta}$ for the general case can be properly defined through a similar process.

2.4. MODEL ASSESSMENT

The proposed approach to construct the probabilistic soil identification framework is based on the Bayesian notion of probability. The proposed approach is capable of incorporating all types of available information, including mathematical models,

field/laboratory data pairs, and subjective engineering experience and judgment. In the Bayesian approach, the unknown parameters $\boldsymbol{\Theta}$ are estimated by the use of the updating rule (Box and Tiao, 1992)

$$p''(\boldsymbol{\Theta}) = \kappa L(\boldsymbol{\Theta})\, p'(\boldsymbol{\Theta}) \tag{2.9}$$

where $p'(\boldsymbol{\Theta})$ is the prior distribution reflecting the state of knowledge about $\boldsymbol{\Theta}$ prior to obtaining the field or experimental data and based on empirical soil classification charts that are currently available, engineering judgment or experience prior to obtaining the set of observations; $L(\boldsymbol{\Theta})$ is the likelihood function representing the objective information on $\boldsymbol{\Theta}$ contained in the field/laboratory data pairs, $\kappa = [\int L(\boldsymbol{\Theta}) p'(\boldsymbol{\Theta}) d\boldsymbol{\Theta}]^{-1}$ is a normalizing factor, and $p''(\boldsymbol{\Theta})$ is the posterior distribution representing the updated state of knowledge about $\boldsymbol{\Theta}$. According to the proposed Bayesian approach, the posterior estimates of $\boldsymbol{\Theta}$ represent the updated state of knowledge about the soil identification that incorporates both the prior information and the objective information from the field/laboratory data pairs.

2.4.1. Likelihood Function

The likelihood function is proportional to the conditional probability of observing the field/laboratory data pairs for a given value of $\boldsymbol{\Theta}$. Assuming that the data are statistically independent, the likelihood function can be expressed as

$$L(\boldsymbol{\Theta}) = \prod_{i=1}^{n} \pi_{\alpha_i|\beta_i} \tag{2.10}$$

where *i* ranges over the n field/laboratory data pairs.

In Eq. (2.10), $\pi_{\gamma|\beta}$ are unknown as in the case with exact measurements, while $\pi_{1\alpha|(\gamma,\beta)}$ comes from calibration and/or engineering judgment about the accuracy of the laboratory tests. In consideration of measurement errors, the likelihood function is then written as

$$L(\boldsymbol{\Theta}) = \prod_{i=1}^{n} \left[\pi_{1\alpha|\beta} \right]_i = \prod_{i=1}^{n} \left\{ \sum_{\gamma} \left[\pi_{1\alpha|(\gamma,\beta)} \cdot \pi_{\gamma|\beta} \right] \right\} \tag{2.11}$$

The Bayesian framework described in the previous section can be used in case of potential laboratory misclassifications, simply using the newly defined $L(\boldsymbol{\Theta})$ in Eq. (2.11).

2.4.2. Prior Distribution

Before accounting for the prior information, the unknown model parameters $\boldsymbol{\Theta}$ are estimated based only on the field/laboratory data by maximizing the likelihood function. Following Richards (1961) the covariance matrix of $\boldsymbol{\Theta}$ is estimated as the negative of the inverse of the Hessian of the log-likelihood functions, $-[\nabla\nabla \ln L(\boldsymbol{\Theta})]^{-1}$ estimated at the maximum-likelihood estimates.

In reality, information about the unknown parameters $\boldsymbol{\Theta}$ is typically available prior to collecting the field/laboratory data pairs. To incorporate such information, the Bayesian approach requires that such prior information is included in the form of the prior distribution of the unknown parameters. The next section provides a specific example of how to construct the prior distribution of $\boldsymbol{\Theta}$.

2.5. APPLICATIONS OF PROBABILISTIC SOIL CLASSIFICATION

In this section, the proposed approach is used to develop two site-specific soil classification charts using data from the TTS near Venice (Italy) and the NGES at Texas A&M University. The applicability of the site-specific chart for other sites in Venice Lagoon is assessed using data from the Malamocco test site, approximately 20 km from TTS.

2.5.1. Application to TTS

As an application of the proposed approach, a site-specific database of field/laboratory data pairs from the TTS (Simonini, 2004) is used. All soils at the site can be grouped into 3 broad categories: medium to fine sands (SP-SM), silt (ML), and clay (CL). They also belong to Classes 3 to 6 in Robertson's classification chart, as shown in Fig. 2.3 (a). The markers used for the data points indicate the classification of each specimen as determined in the laboratory. Areas 5 and 6, grouping all sand mixtures, represent sand mixtures (here called area A), area 4 represents silt mixtures (area B), and area 3 indicates clays (area C). Consequently, areas 1, 8, and 9 are eliminated and it is decided to combine areas 5 and 6 and have only 3 areas and two boundaries. Fig. 2.2 shows the simplified chart in the semi-logarithmic space.

The probabilistic identification assuming exact measurements is considered first. As described in a previous section, the vector of unknown model parameters is $\boldsymbol{\Theta} = (a_1, m_1, \theta_a, \theta_m, \pi_{1|A}, \ \pi_{1|B}, \pi_{1|C}, \pi_{2|A}, \pi_{2|B}, \pi_{2|C})$. Tables 2.4 and 2.5 list the maximum

likelihood estimates (MLE) of $\boldsymbol{\Theta}$, separating the boundary variables from the conditional probabilities. The values of the parameters are obtained by maximizing the likelihood function in Eq. (2.8). Given the scarcity of the data used to assess the model parameters, other choices of a_1, m_1, θ_a, and θ_m are also possible, if they also provide the same value of the likelihood function. Due to the lack of uniqueness of the maximum-likelihood estimates, the estimates of the standard deviations and the correlation coefficients for a_1, m_1, θ_a, and θ_m obtained using the negative of the inverse of the Hessian of the log-likelihood function evaluated at the maximum-likelihood estimate (Richards, 1961) are not accurate. Therefore, this work provides only possible estimates of a_1, m_1, θ_a, and θ_m that maximise the likelihood function, but not the corresponding standard deviation and correlation coefficients.

Table 2.4. Point estimates of the unknown boundary parameters with exact measurements for TTS

Method	a_1	m_1	θ_a	θ_m
Robertson's chart	1.260	0.457	0.787	0.172
MLE	0.479	0.925	1.529	0.243
Posterior (COV=0.1)	1.259	0.451	0.776	0.169
Posterior (COV=0.6)	0.608	0.892	1.719	0.127

Table 2.5. Maximum likelihood estimates of the unknown parameters with exact measurements for TTS

	$\pi_{1\|A}$	$\pi_{1\|B}$	$\pi_{1\|C}$	$\pi_{2\|A}$	$\pi_{2\|B}$	$\pi_{2\|C}$
Mean	0.474	0.081	0.053	0.188	0.717	0.132
Standard deviation	1.029	0.905	0.036	0.378	0.998	0.055
Correlation coefficients						
$\pi_{1\|B}$	0.062					
$\pi_{1\|C}$	–0.073	–0.006				
$\pi_{2\|A}$	–0.984	–0.063	0.071			
$\pi_{2\|B}$	–0.03	–0.102	0.013	0.028		
$\pi_{2\|C}$	0.024	–0.002	–0.089	–0.024	–0.009	

The Bayesian approach requires information on the parameters **Θ** available before collecting the field/laboratory data pairs to be included as a prior distribution. In this study, the distribution models for the prior distribution are selected looking at the range of each parameter. The means are based on the conventional classification chart (Robertson, 1990), and the standard deviations are based on an assumed value for the coefficient of variation (COV).

Table 2.6. Prior distributions, means, and standard deviations for TTS

Parameter ranges	Distribution models	Mean	Standard deviation	
			COV=0.1	COV=0.6
$-\infty \leq a_1 \leq \infty$	Normal	1.260	0.127	0.759
$0 \leq m_1 \leq \infty$	Log-Normal	0.457	0.046	0.274
$0 \leq \theta_a \leq \infty$	Log-Normal	0.787	0.079	0.472
$0 \leq \theta_m \leq \infty$	Log-Normal	0.172	0.017	0.103
$0 \leq \pi_{1\|A} \leq 1$	Beta	0.700	0.070	0.420
$0 \leq \pi_{1\|B} \leq 1$	Beta	0.150	0.015	0.090
$0 \leq \pi_{1\|C} \leq 1$	Beta	0.100	0.010	0.060
$0 \leq \pi_{2\|A} \leq 1$	Beta	0.200	0.020	0.120
$0 \leq \pi_{2\|B} \leq 1$	Beta	0.700	0.070	0.420
$0 \leq \pi_{2\|C} \leq 1$	Beta	0.200	0.020	0.120

Two cases are considered to explore the effect of the prior distribution on the posterior estimates of **Θ,** COV=0.1 and COV=0.6. The COV reflects the confidence in the accuracy of the prior estimates. So COV=0.1 means a relatively high confidence in the validity of Robertson's chart, while COV=0.6 puts little weight on the prior estimates. Table 2.6 lists the assumed prior distribution models, means and standard deviations for each of the unknown parameters. Tables 2.4, 2.7 and 2.8 list the posterior

statistics of the parameters $\boldsymbol{\Theta}$ for COV=0.1 and COV=0.6. An algorithm for computing these statistics based on importance sampling is described in Gardoni (2002).

Table 2.7. Posterior statistics of the unknown parameters with exact measurements for COV=0.1 (TTS)

	$\pi_{1\|A}$	$\pi_{1\|B}$	$\pi_{1\|C}$	$\pi_{2\|A}$	$\pi_{2\|B}$	$\pi_{2\|C}$
Mean	0.701	0.149	0.099	0.199	0.703	0.198
Standard deviation	0.006	0.004	0.003	0.005	0.006	0.005
Correlation coefficients						
$\pi_{1\|B}$	0.001					
$\pi_{1\|C}$	–0.003	0.004				
$\pi_{2\|A}$	–0.885	–0.003	0.001			
$\pi_{2\|B}$	–0.004	–0.694	0.001	0.002		
$\pi_{2\|C}$	0.002	–0.003	–0.039	0.001	0.001	

Table 2.8. Posterior statistics of the unknown parameters with exact measurements for COV=0.6 (TTS)

	$\pi_{1\|A}$	$\pi_{1\|B}$	$\pi_{1\|C}$	$\pi_{2\|A}$	$\pi_{2\|B}$	$\pi_{2\|C}$
Mean	0.604	0.074	0.053	0.307	0.833	0.104
Standard deviation	0.835	0.959	0.065	0.823	1.119	1.048
Correlation coefficients						
$\pi_{1\|B}$	0.164					
$\pi_{1\|C}$	0.017	–0.256				
$\pi_{2\|A}$	0.066	0.126	–0.007			
$\pi_{2\|B}$	–0.164	–0.999	0.256	–0.126		
$\pi_{2\|C}$	0.131	0.038	–0.079	0.146	–0.038	

Fig. 2.5 shows the results of the soil identification using (a) MLE estimates, (b) posterior means based on COV=0.1, and (c) posterior means based on COV=0.6. The dashed lines are the boundary lines based on Robertson (1990), the double dashed lines are the boundary lines obtained using the MLE estimates, and the thin solid lines are the boundary lines obtained using the posterior means.

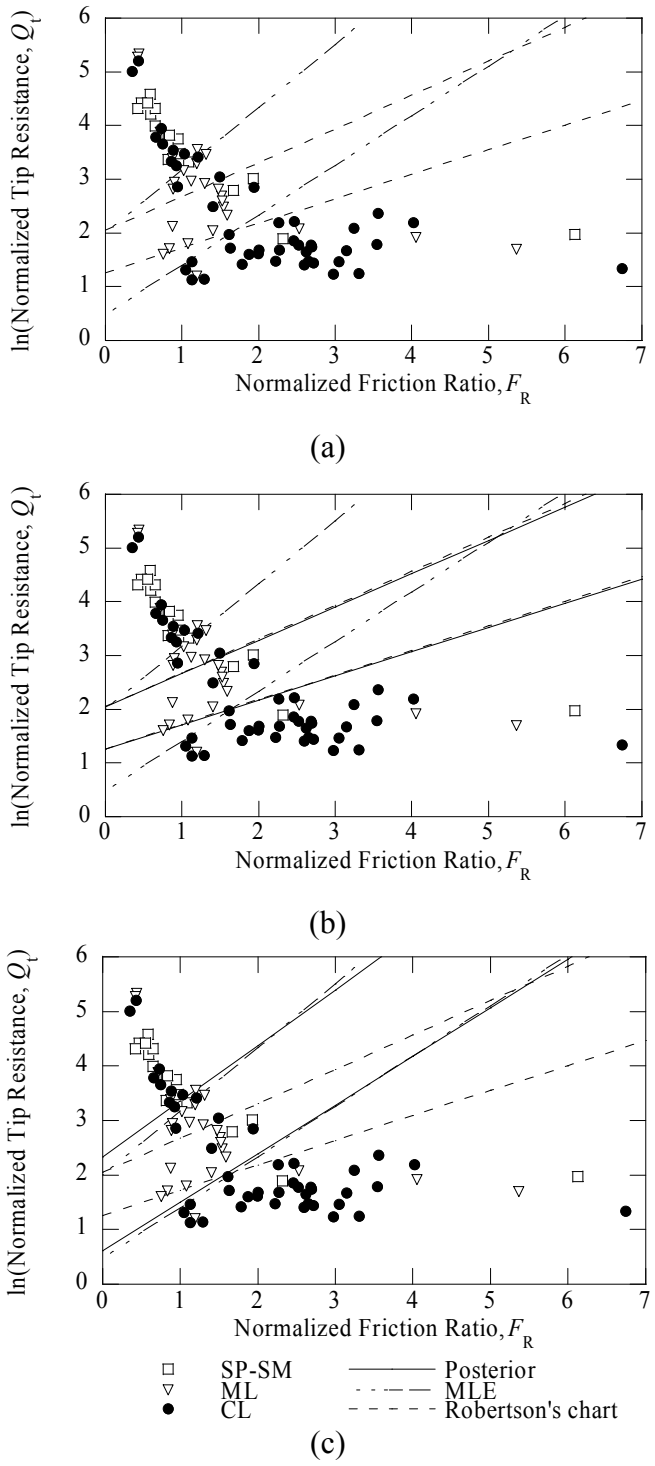


Fig. 2.5. Results of soil identification estimated by Bayesian method with exact measurements (a) MLE (b) posterior estimate with COV=0.1 (c) posterior estimate with COV=0.6. (TTS)

If, based on engineering judgment, COV=0.6 is believed to be more appropriate (i.e., Table 2.8), then the estimates of $\pi_{\gamma|\beta}$ indicate that for a field experiment with values of F_R and Q_t in area A, the soil has 60.4% probability of being a sand mixture (SP-SM), 30.7% probability of being a silt mixture (ML), and 8.9% probability of being a clay mixture (CL). The results can give similar estimates of probability for each soil class if a field experiment gives values of F_R and Q_t in area B or C.

The following observations can be made from these results:

1. When a small COV (e.g., 0.1) is used, the predictions are almost the same as in the original classification chart based on Robertson (1990) (Fig. 2.5(b).) This is because a high level of confidence is placed on the prior statistics of the parameters $\boldsymbol{\Theta}$.
2. When a larger COV (0.6) is used, the prior information has a limited effect on the posterior statistics, which are close to the maximum likelihood statistics (Fig. 2.5(c).)
3. The soil identification based on posterior and maximum likelihood estimates of the boundary parameters provide a more accurate classification of the soil samples in the database.

The presence of errors in the laboratory classification of soils may also influence the location of the boundary lines and the probability of a certain CPT measurement to indicate the correct soil. To incorporate the effect of imprecise measurements in the example presented here, the laboratory tests are assumed to be correct in identifying sands (see Table 2.9). This is a reasonable assumption since sieve analysis procedures

and interpretations are easy to follow and are unlikely to be performed incorrectly. On the other hand, determination of Atterberg limits is largely influenced by the tester's perception and more likely to lead to incorrect classification, especially with the low plasticity silt and clay mixtures at TTS. For these reasons, the conditional probabilities of incorrect classification in Table 2.9 are much larger for the fine-grained soils. The specific values selected for this example are exclusively based on engineering judgment. If the specimens were tested again, actual estimates of the error may be obtained. This example is intended to show how imprecise measurements can be incorporated into the analysis. A parametric study may also be performed to evaluate the influence of different scenarios.

The MLE of $\boldsymbol{\Theta}$ including measurement errors are listed in Tables 2.10 and 2.11. The values of the parameters are maximised with the likelihood function in Eq. (2.11).

Table 2.9. Conditional probabilities based on laboratory test, soil class, and areas for TTS

Laboratory test ($I\alpha$)	Areas (β)		
	A	B	C
	Soil class (γ)=sand		
Sand	1	1	1
Silt	0	0	0
Clay	0	0	0
	Soil class (γ)=silt		
Sand	0	0	0
Silt	0.7	0.8	0.6
Clay	0.3	0.2	0.4
	Soil class (γ)=clay		
Sand	0	0	0
Silt	0.3	0.4	0.2
Clay	0.7	0.6	0.8

Table 2.10. Point estimates of the unknown boundary parameters with measurement errors (TTS)

Method	a_1	m_1	θ_a	θ_m
Robertson's chart	1.260	0.457	0.787	0.172
MLE	0.658	0.826	1.522	0.229
Posterior (COV=0.1)	1.089	0.418	0.719	0.165
Posterior (COV=0.6)	1.049	0.569	1.689	0.159

Table 2.11. Maximum likelihood estimates of the unknown parameters with measurement errors (TTS)

	$\pi_{1\|A}$	$\pi_{1\|B}$	$\pi_{1\|C}$	$\pi_{2\|A}$	$\pi_{2\|B}$	$\pi_{2\|C}$
Mean	0.494	0.130	0.107	0.223	0.783	0.159
Standard deviation	0.963	0.756	0.879	1.009	0.953	1.009
Correlation coefficients						
$\pi_{1\|B}$	–0.344					
$\pi_{1\|C}$	–0.177	–0.492				
$\pi_{2\|A}$	–0.026	–0.265	–0.115			
$\pi_{2\|B}$	–0.09	–0.361	–0.191	–0.036		
$\pi_{2\|C}$	–0.022	–0.260	–0.111	0.025	–0.033	

Table 2.12. Posterior statistics of the unknown parameters with measurement errors for COV=0.1 (TTS)

	$\pi_{1\|A}$	$\pi_{1\|B}$	$\pi_{1\|C}$	$\pi_{2\|A}$	$\pi_{2\|B}$	$\pi_{2\|C}$
Mean	0.676	0.145	0.097	0.219	0.706	0.194
Standard deviation	0.004	0.008	0.009	0.009	0.001	0.012
Correlation coefficients						
$\pi_{1\|B}$	0.999					
$\pi_{1\|C}$	0.215	0.215				
$\pi_{2\|A}$	–0.273	–0.273	–0.117			
$\pi_{2\|B}$	–0.999	–0.999	–0.215	0.273		
$\pi_{2\|C}$	–0.759	–0.759	0.227	0.676	0.759	

Table 2.13. Posterior statistics of the unknown parameters with measurement errors for COV=0.6 (TTS)

	$\pi_{1\|A}$	$\pi_{1\|B}$	$\pi_{1\|C}$	$\pi_{2\|A}$	$\pi_{2\|B}$	$\pi_{2\|C}$
Mean	0.675	0.121	0.076	0.212	0.775	0.236
Standard deviation	0.924	0.843	0.804	0.781	0.992	1.079
Correlation coefficients						
$\pi_{1\|B}$	–0.279					
$\pi_{1\|C}$	–0.335	–0.473				
$\pi_{2\|A}$	–0.996	0.286	0.325			
$\pi_{2\|B}$	0.278	–0.995	0.463	–0.285		
$\pi_{2\|C}$	0.009	–0.067	–0.108	–0.009	0.068	

Similarly, the unknown parameters derived from the posterior statistics are listed in Tables 2.10, 2.12 and 2.13. Fig. 2.6 shows a comparison between probabilistic identification results obtained with exact measurements and with measurement errors.

The following observations can be made from these results:

1. For the MLE (Fig. 2.6(a)), the influence of measurement errors is minimal, since the likelihood function depends on the measurements only, irrespective of their accuracy.
2. In the posterior estimates for the case of small COV (e.g., 0.1), the results strongly depend on the SP-SM soil specimens rather than the ML and CL samples. This is a consequence of the assumptions on the measurement errors, which place complete confidence on the laboratory classification of sands but expect errors in the classification of fine-grained soils (Fig. 2.6(b).) The posterior statistic reflects measurement errors as well as the prior information.
3. When a larger COV (e.g., 0.6) is used, the influence of the prior information is limited, as it was when no measurement errors were considered. However, the

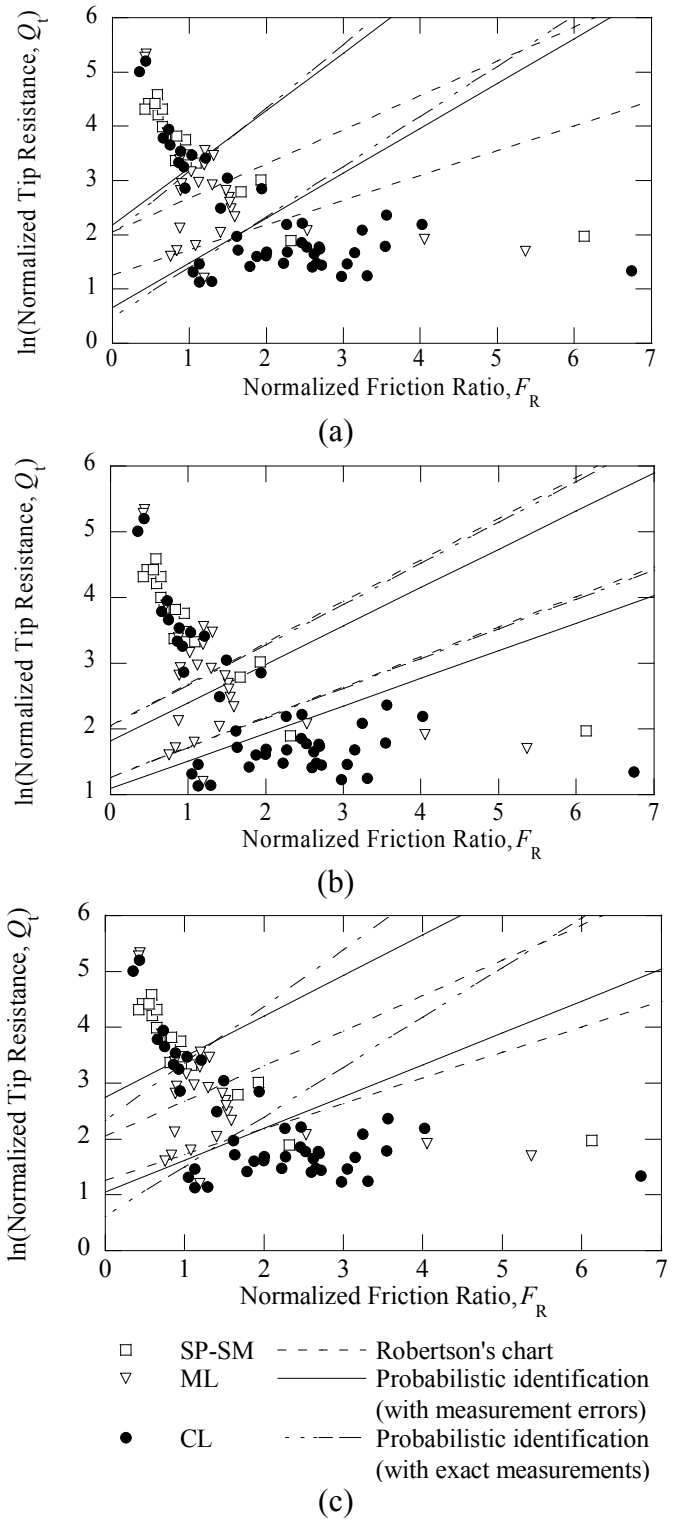


Fig. 2.6. Comparison of results with exact measurements and measurement errors (a) MLE (b) posterior estimate with COV=0.1 (c) posterior estimate with COV=0.6. (TTS)

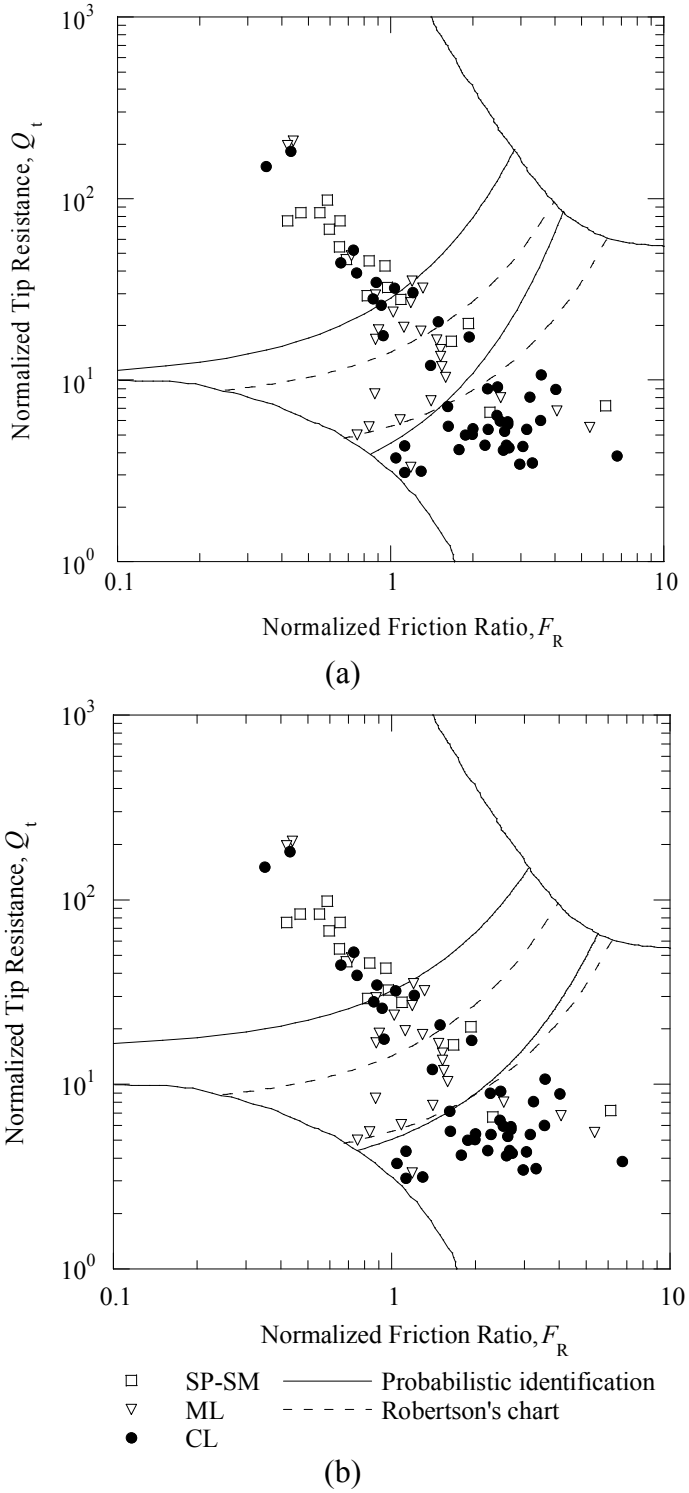


Fig. 2.7. Probabilistic soil identification with TTS CPT data using (a) exact measurements (b) measurement errors

results show that the laboratory data are given less importance than in the case with exact measurements, as evidenced by the larger difference between the posterior estimates and the MLE in this case (Fig. 2.6(c).)

As an example of the effect of measurement errors on Robertson's chart, Fig. 2.7 shows the posterior estimates of the boundary lines for prior estimates with COV=0.6 in the case of exact measurements and with measurement errors.

As described in the introduction, the probabilistic identification chart is meant to improve the identification of soils through CPT on a regional basis. To test the applicability of the developed site-specific chart, additional data not included in the assessment of the chart are considered. The data are from the Malamocco test site (Ricceri et al., 2002), which is still in Venice Lagoon, but approximately 20 km from TTS. Fig. 2.8 shows the clay specimens as full circles, silts as hollow triangles, and sands as hollow squares, as classified by Ricceri et al. through laboratory testing. Considering exact measurements (Fig. 2.8(a)), the identification based on the developed probabilistic site-specific chart is considerably more accurate than that based on the original Robertson's chart, especially for the silt samples (Table 2.14). Note that in the original data set no samples fell in the upper right corner of the Robertson's chart (areas 8 and 9), therefore they were not considered. Fig. 2.8(b) includes measurement errors in the classification of the TTS samples. Again, the identification of silts and sands based on the probabilistic chart is more accurate than the one based on Robertson's chart (Table 2.14). However, the improvement in the identification of clays is reduced and clay samples are still prone to misclassification.

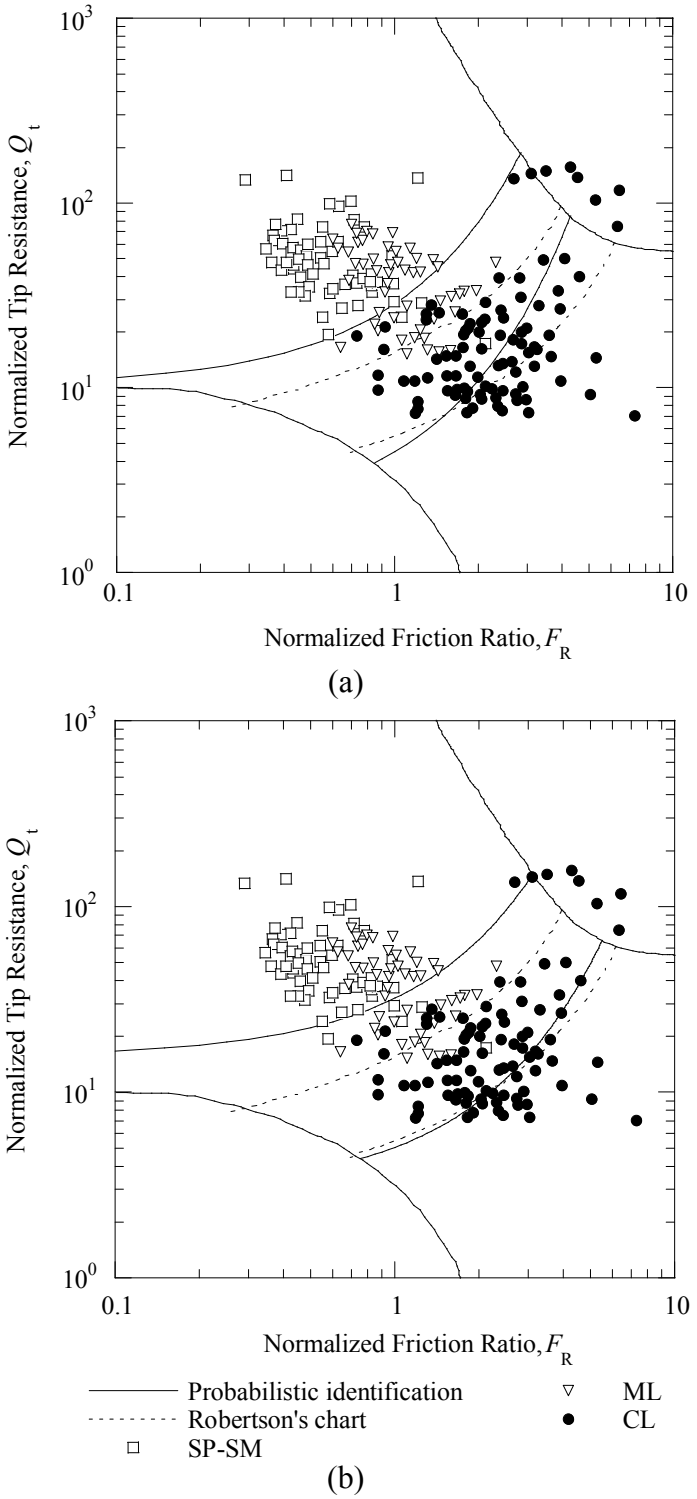


Fig. 2.8. Probabilistic soil identification at the Malamocco Test Site with (a) exact measurements and (b) including measurements errors

While the Malamocco data are used here only for validation purposes, in practice a new updated Venice regional chart would be created including this set of data, and any additional sets, as they become available.

Table 2.14. Comparison between accuracy of Robertson's chart and the developed site-specific chart with exact measurement and with measurement errors for the Malamocco Test Site data

Soil type	Total data points	Number of correct points (Percentage)		
		Robertson's chart	Site-specific chart with exact measurements	Site-specific chart with measurement errors
SP-SM	61	60 (98%)	58 (95%)	56 (92%)
ML	51	4 (8%)	21 (41%)	22 (43%)
CL	92	19 (21%)	39 (42%)	30 (33%)
Total	204	83 (41%)	118 (58%)	108 (53%)

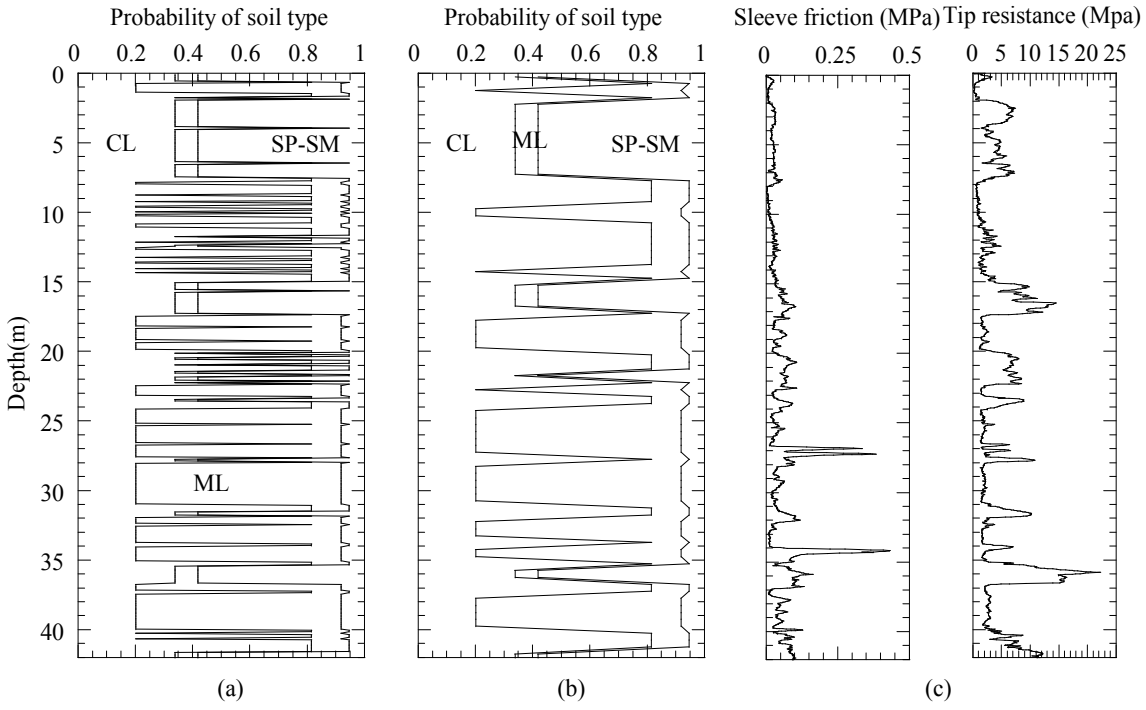


Fig. 2.9. Probability of identifying each soil class with depth at TTS with discretization interval of (a) 10 cm and (b) 50 cm, and (c) CPT records

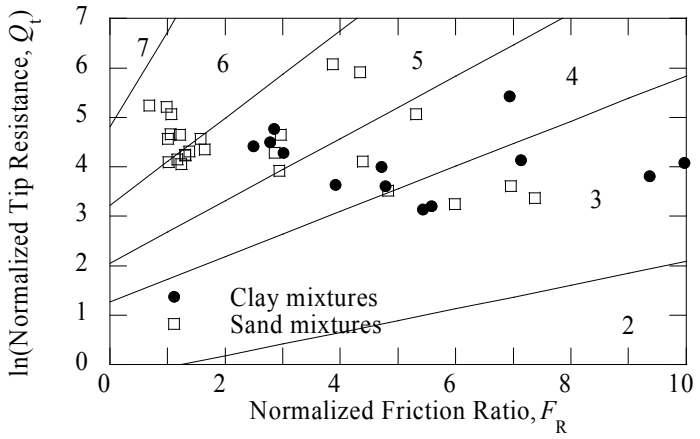
The analytical description of the boundaries and the ability to estimate the probability of soil class given CPT results can also be exploited to create interpretations of CPT logs such as shown in Fig. 2.9 showing the probability of each soil class with depth. The discretization and averaging of the CPT record with depth is an important issue, but the comparison of the profiles with intervals of 10 cm and 50 cm can give an indication of the main soil units, as well as the variability at the site.

2.5.2. Application to NGES

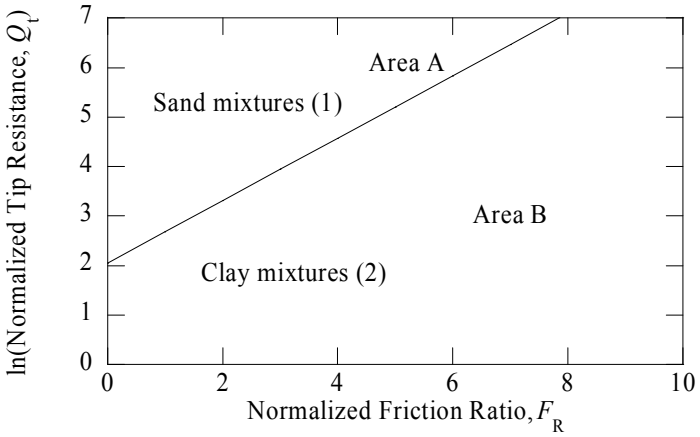
From the late 1970's, numerous in situ and laboratory tests including CPT, grain size distribution, and consistency limits were performed at the NGES located in the Riverside campus of Texas A&M University. All available information for the NGES is summarised by Briaud (1997). For the purpose of soil identification, the field/laboratory data pairs are presented in Fig. 2.10(a) on a transformed chart. The laboratory based classification for each specimen is indicated by the marker used for each data point.

Because of the limited information available for each soil sample, it was decided to divide the soils into only two possible classes, clay mixtures or sand mixtures. After simplification, the chart is reduced to two areas, A and B, and one boundary between them as shown in Fig. 2.10(b).

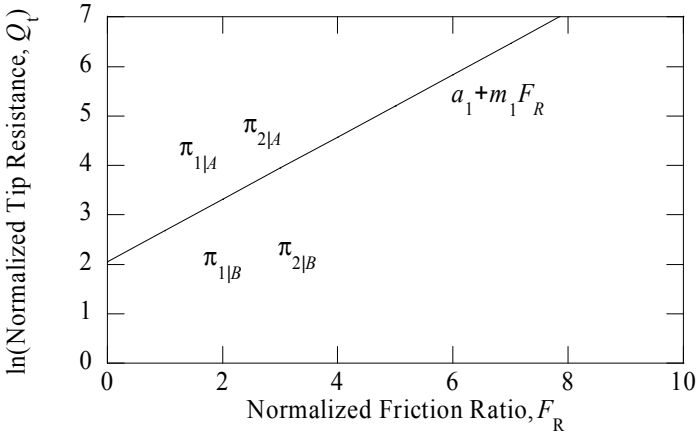
Fig. 2.10(c) shows the six unknown model parameters, which comprise four conditional probabilities and two boundary parameters. After applying possible constraints (Eq. (2.4)), only four unknown independent model parameters $\boldsymbol{\Theta} = (a_1, m_1, \pi_{1|A}, \pi_{1|B})$ remain.



(a)



(b)



(c)

Fig. 2.10. (a) Transformed Robertson's classification chart with NGES data (b) simplified chart (c) sets of model parameters

The probabilistic soil identification was modelled first assuming exact measurements and then including measurement errors, as in the previous example. The prior information for the unknown parameters is listed in Table 2.15. Only the point estimates for the unknown boundary parameters are provided. As previously discussed, standard deviations and correlation coefficient are not accurate because of the lack of uniqueness in the MLE estimates.

Table 2.15. Prior distributions, means, and standard deviations for NGES

Parameter ranges	Distribution models	Mean	Standard deviation	
			COV=0.1	COV=0.6
$-\infty \le a_1 \le \infty$	Normal	2.053	0.205	1.232
$0 \le m_1 \le \infty$	Log-Normal	0.629	0.063	0.378
$0 \le \pi_{1\|A} \le 1$	Beta	0.700	0.070	0.420
$0 \le \pi_{1\|B} \le 1$	Beta	0.150	0.015	0.090

Table 2.16 and Fig. 2.11 show the results of the analysis. Similarly to what was observed for TTS, with the small value of COV (0.1), the posterior estimate is close to Robertson's original chart, while the posterior estimate with a larger COV (0.6) is close to the MLE.

Table 2.16. Point estimates of the unknown boundary parameters with exact measurement for NGES

Method	a_1	m_1
Robertson's chart	2.053	0.629
MLE	1.672	0.549
Posterior (COV=0.1)	2.052	0.621
Posterior (COV=0.6)	1.856	0.515

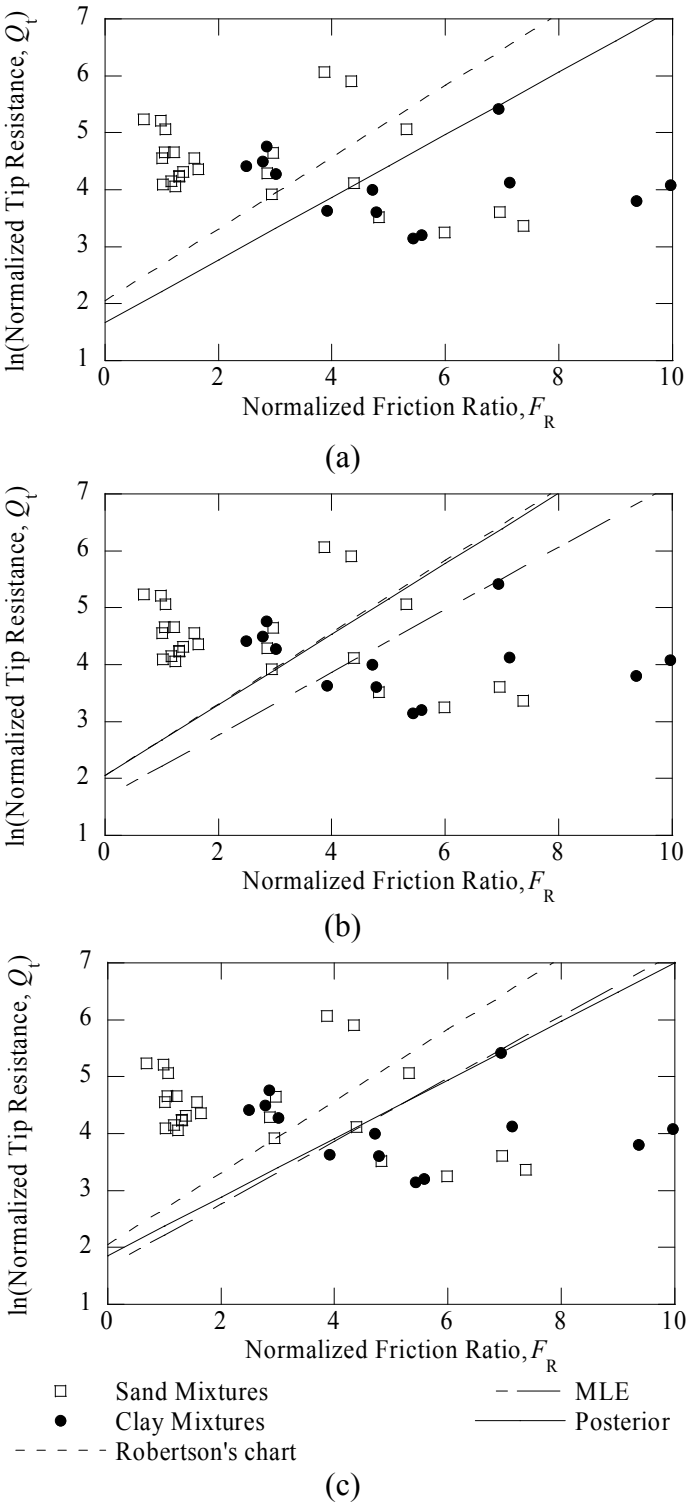


Fig. 2.11. Results of soil identification estimated by Bayesian method with exact measurements (a) MLE (b) posterior estimate with COV=0.1 (c) posterior estimate with COV=0.6. (NGES)

Table 2.17 lists the conditional probabilities that a laboratory test will indicate a certain soil class correctly and incorrectly. The values are again selected based on judgment, rather than statistical measurements. Fig. 2.12 shows the effect of incorporating measurement errors while Table 18 lists the unknown boundary parameters. The posterior estimates in the case of COV=0.6 (i.e., low confidence on Robertson's chart) and MLE are basically the same, irrespective of the inclusion of measurement errors. This happens because in the MLE the data control the position of the boundary and any further lowering of the boundary to classify correctly an additional sand sample would result in the wrong classification for a larger number of clay

Table 2.17. Conditional probabilities based on laboratory test, soil class, and area for NGES

Laboratory test ($I\alpha$)	Areas (β)	
	A	B
	Soil class (γ)=sand mixtures	
Sand mixtures	1	1
Clay mixtures	0	0
	Soil class (γ)=clay mixtures	
Sand mixtures	0.3	0.2
Clay mixtures	0.7	0.8

Table 2.18. Means of the unknown boundary parameters with measurement errors for NGES

Method	a_1	m_1
Robertson's chart	2.053	0.629
MLE	1.599	0.565
Posterior (COV=0.1)	1.966	0.582
Posterior (COV=0.6)	1.842	0.517

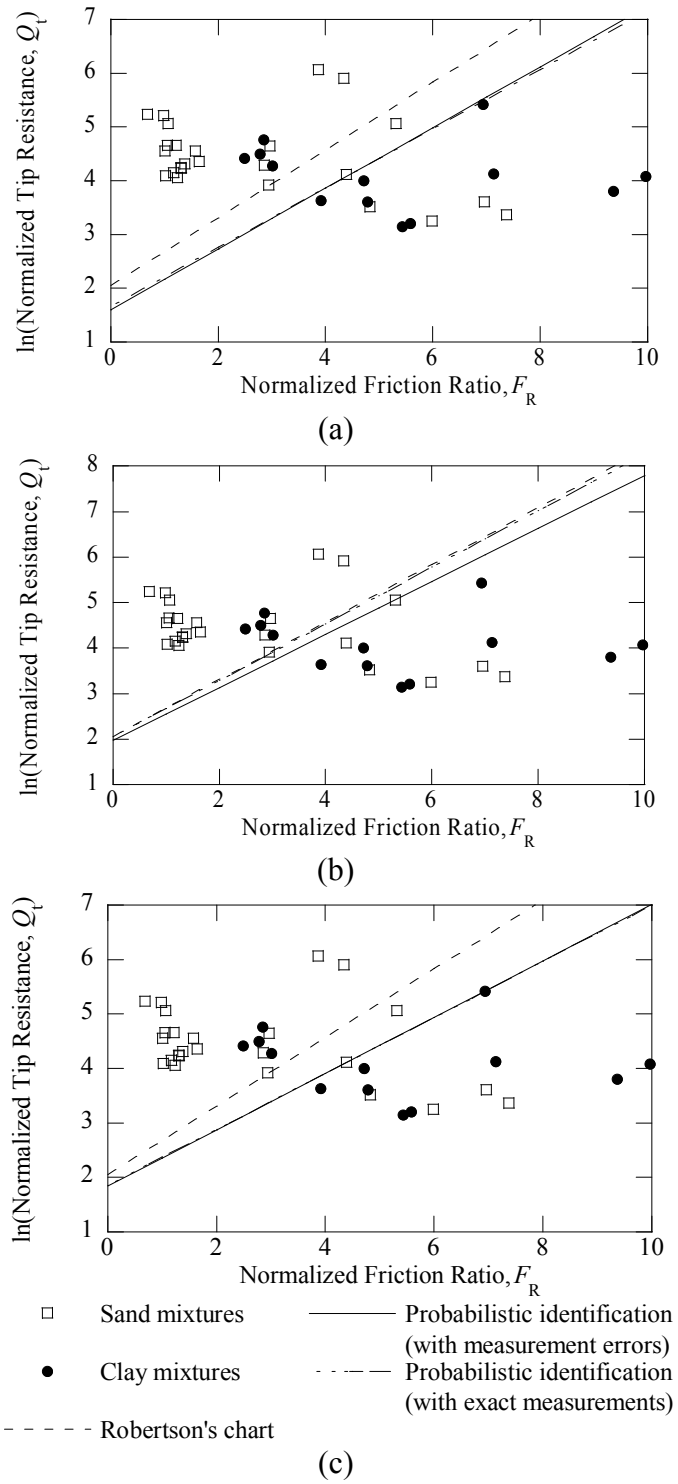


Fig. 2.12. Comparison of results with exact measurements and measurement errors (a) MLE (b) posterior estimation with COV=0.1 (c) posterior estimation with COV=0.6. (NGES)

specimens, irrespective of laboratory classification errors. The larger COV also assigns low importance to the prior estimates and allows data to control the posterior estimates.

Conversely, if COV=0.1, the prior estimates control the position of the boundary, but the introduction of measurement errors places some importance on the correct sand classification of one additional specimen, rather than the perhaps incorrect classification of the clay samples in Fig. 2.12(b), moving the boundary slightly lower than it was in the case with no consideration for measurement errors.

The application of a sensitivity analysis is suggested when judgment is used to quantify measurement errors in the application of the proposed model. This is especially worthwhile in situations where considerable variation in laboratory measurement is expected. A different conditional probability may also be assigned to a certain sub-group or even each of the soil samples. For example, the 4 clay specimens identified by a marker shaped as a cross in Fig. 2.13 are assigned a 50% probability of being classified either as a sand or clay mixture by laboratory classification irrespective of the area in which they fall. All other specimens are assigned the same probabilities of measurement errors as in the earlier case (Table 2.17).

Fig. 2.13 shows the sensitivity of the MLE to measurement errors. The new boundary is determined by the sand mixtures, as a consequence of the assumption made on the clay specimens. Since the 4 marked samples at the boundary could be sand or clay mixtures, the boundary line takes the lowest possible position that increases the number of correct classifications of sand specimens, without increasing the number of incorrect classifications of clays.

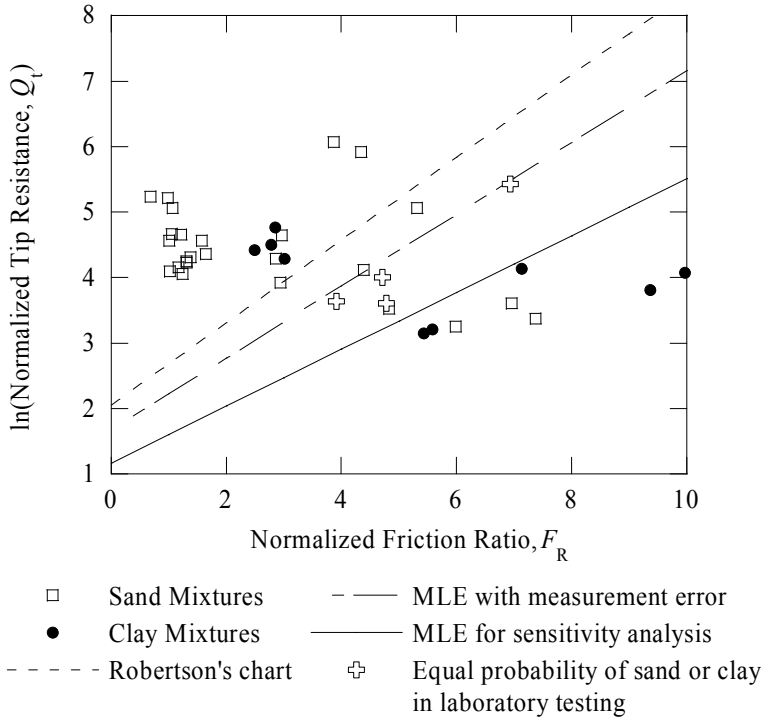


Fig. 2.13. Result of soil identification sensitivity by MLE with NGES CPT data

2.6. CONCLUSIONS

Soil classification charts are used to identify soil types and subsurface profile at a site. Current soil classification charts based on CPT data give a general soil class interpretation, but they do not reflect the site-specific information that is typically available.

A probabilistic approach is proposed to develop site-specific soil identification based on CPT measurements. The proposed approach provides a simple and straightforward tool that allows updating the soil classification charts based on site-specific data. The probability that a soil is correctly classified is also estimated. The site-specific classification chart can be used for a more accurate identification of the soil

and accounts both for prior information available before conducting the tests and for the site-specific data. As an illustration, the proposed approach is implemented using CPT data from the Treporti Test Site near Venice (Italy) and NGES at Texas A&M University to develop two site-specific charts. The applicability of the first chart for other sites in Venice Lagoon is assessed using data from the Malamocco test site that is approximately 20 km from the Treporti Test Site.

In practice, the developed approach can be used to construct and update regional classification charts that are be more accurate than a generic chart within the geographical area of interest. In addition, as illustrated in an example, the information can be used to produce interpretations of CPT logs, giving a profile of the probability of identifying each of the soil classes with depth.

CHAPTER III

BAYESIAN UPDATING OF A UNIFIED COMPRESSION MODEL

3.1. INTRODUCTION

Any structure is subjected to some amount of differential settlements due to the natural heterogeneity of the soil, even in the most uniform deposit. Usually, highly compressible soils are viewed as more likely to result in significant differential settlements simply because the average settlement is high. However, soils characterized by low or medium compressibility can still lead to excessive distortion in the structure if the soil deposit is highly variable. Two different issues contribute to the problem: (1) the assumed soil profile may differ from location to location at the site; and (2) the material properties and initial conditions within one layer may vary due to changes in composition, gradation, or natural water content. In order to properly assess the reliability of a structure with respect to differential settlements, it is necessary to characterize both the aleatory and epistemic uncertainties associated with the natural variability of the soil and with the model used to describe the physical phenomenon.

Traditionally, the compression response of saturated clays and prevalently clayey soils has been described by a bi-linear curve in the logarithm of the vertical effective stress versus the void ratio ($e - \log p'_v$) space. The compression index, the swelling index, and the preconsolidation pressure, which characterize the consolidation curve, are functions of the soil composition, initial void ratio, stress conditions, and stress history to which the soil was subjected in the field.

The settlements of sands are usually estimated using elasticity theory, often corrected with empirically based coefficients. Silts and any type of mixtures are considered to behave as "clays" or "sands" depending on the compressibility of the material.

While this piecemeal approach is adequate when estimating the one-dimensional settlement of a given deterministic soil profile, its value for the reliability analysis of differential settlements is severely limited because of the difficulty in addressing the two issues outlined above. Biscontin et al. (2007) developed a unified compression model which is able to describe the response of saturated soils in a wide range of stresses, composition, and initial conditions. In this work, we develop a probabilistic compression model that properly accounts for the prevailing uncertainties, including model errors and statistical uncertainty. We also present a Bayesian formulation that can be applied to update the parameters of the current probabilistic model with new laboratory observations as they become available. In the proposed probabilistic model, any bias in the deterministic estimates of the unified compression model parameters is corrected. A probabilistic framework provides more accurate estimation of settlements to improve the prediction of differential settlement and helps in the selection of mitigation and intervention strategies.

This chapter is composed of four sections. Following this introduction, the unified soil compression model and a set of laboratory test results are described. Then, a probabilistic model and a Bayesian updating approach are developed. The final section presents an application of this study.

3.2. DETERMINISTIC MODEL

3.2.1. Unified Soil Compression Model

Biscontin et al. (2007) present a unified soil compression model specifically developed to describe the compression response of soils that can be characterized by the combination of cohesive and granular materials. The model assumes that the one-dimensional compression response converges to a linear limiting compression curve (K_0-LCC) in a double logarithmic space $(\log e - \log p'_v)$ at high stress level as shown in Fig. 3.l. The K_0-LCC is described by its slope ρ_c, and a reference void ratio e_{1v}, at the atmospheric pressure p_{at}

$$\log\left(\frac{e}{e_{1v}}\right) = -\rho_c \log\left(\frac{p'_v}{p_{at}}\right) \tag{3.1}$$

where e = the current void ratio, p'_v = the current effective stress.

Soil is idealized as a combination of a compressible matrix (the clay-water phase) and incompressible inclusions (sand and silt particles). Assuming that ρ_c is the same irrespective of grain size for a particular soil mineralogy, the position of the K_0-LCC, defined by the reference void ratio e_{1v}, becomes a function of the relative abundance of the clay-water phase on the granular phase and the reference void ratios of the two phases only:

$$e_{1v} = e_{g1v}(1 - FF) - FF \qquad FF \le 0.2 \tag{3.2}$$

$$e_{1v} = e_{g1v}\exp(0.25 - 4.76FF) + (e_{c1v} - 0.12)FF \qquad 0.2 < FF < 0.7 \tag{3.3}$$

$$e_{1v} = e_{c1v}FF \qquad FF \ge 0.7 \tag{3.4}$$

where FF = gravimetric clay fraction, e_{g1v} = reference void ratio for the K_0-LCC of the granular phase only, and e_{c1v} = reference void ratio for the K_0-LCC of the clay-water phase only.

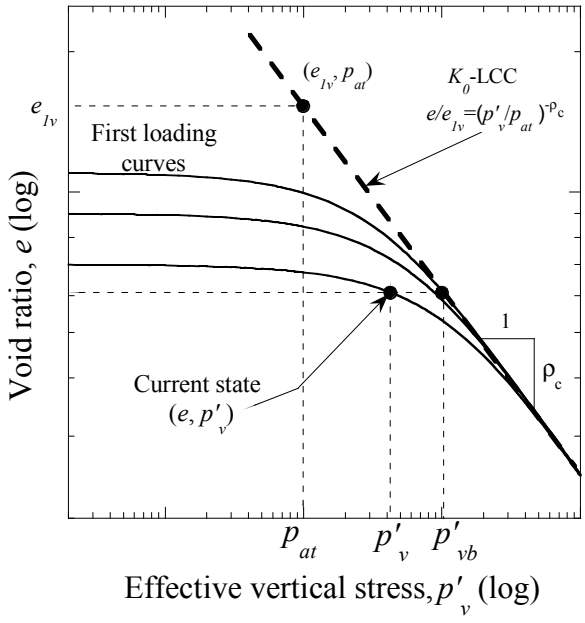


Fig. 3.1. Idealized 1D compression response of soils with different initial density (adapted from Biscontin et al. 2007)

Disregarding the contribution from elastic strains, the compression response from the initial conditions to the K_0-LCC is described by the following relationship:

$$\frac{de}{e} = -\rho_c \left\{ 1 - \left[1 - \frac{p'_v}{p_{at}} \left(\frac{e}{e_{1v}} \right)^{1/\rho_c} \right]^{\alpha} \right\} \frac{dp'_v}{p'_v} \tag{3.5}$$

where α = model parameter controlling the curvature of the compression curve in the transitional regime preceding the K_0-LCC.

This model presents the advantage of requiring few input parameters (ρ_c, e_{c1v} and e_{g1v}) that are uniquely characterized by the mineralogy of the soils, but are also able to describe a wide range of responses. Parameter α seems to depend mostly on the gradation characteristics of the soils, rather than mineralogy. At a site, each layer comprised of soils with relatively similar mineralogy but varying grain size distributions can be assigned a set of these parameters. The variation in soil response within one layer can then be captured by a varying clay fraction *FF* with both depth and location at the site.

3.2.2. Laboratory Database

Laboratory one-dimensional compression tests are used to construct the probabilistic model to estimate the statistics of the deterministic model parameters, as well as the additive model error. The database presented by Biscontin et al. (2007) is used to illustrate the probabilistic model.

The Biscontin et al.'s database consists of three sets of one-dimensional laboratory tests on natural samples, artificially reconstructed mixtures, and reconstituted samples. The natural samples belong to all three soil classes present at the site, silty sands (SP-SM), silts (ML), and clays (CL). The artificially reconstructed mixtures were prepared at different clay fractions and loaded to high stress levels to investigate the effect of *FF* on e_{c1v}. The reconstituted soil specimens were simply some of the natural SP-SM and ML samples tested again at higher stress levels to ensure the K_0-LCC regime was actually reached.

The first series of data is composed of 30 laboratory compression tests, 18 on CL, 7 on ML, and 5 on SP-SM. The majority of the specimens was loaded up to a maximum vertical stress of 6 MPa, with a few loaded up to 12 MPa. Among these tests, only 9 clearly reached the K_0-LCC regime and ρ_c. We note that each one-dimensional compression curve has a varying number of data points n_k where k indicates a specific sample. Therefore, the database was compiled from 100 data points on the natural samples, 60 on the artificially reconstructed mixtures, and 238 on the reconstituted samples. Table 3.1 lists initial void ratio e_0 and FF measured in the laboratory for all samples used in this work.

3.2.3. Deterministic Estimates

The soils in the database presented above are characterized by a relatively uniform mineralogy across all gradation sizes, and therefore can be described by a single set of parameters. Biscontin et al. (2007) estimated values of 1.94 for e_{g1v}, 2.0 for e_{c1v}, 0.24 for ρ_c, and 0.65 for α based on visual inspection of the results. Fig. 3.2 shows the deterministic estimates of the void ratio versus the measured ones. The laboratory data points are shown as triangles (△) for the natural samples, circles (○) for the artificially reconstructed mixtures, and squares (□) for the reconstituted samples. For a perfect model, the data should line up along the 1:1 solid line. As shown in Fig. 3.2, the deterministic estimates for all samples show the same bias. The deterministic estimates

also show more scatter for the reconstituted samples than for the natural samples and the artificially reconstructed mixtures.

Table 3.1. Initial void ratio and clay fraction for the one-dimensional compression tests

Sample	Type of soil	e_0	*FF*
Natural Samples			
MSM10-6	CL	0.903	0.60
MSM10-14	CL	1.091	0.78
MSM10-28	CL	0.667	0.41
MSM10-43	CL	0.675	0.60
MSM10-48	CL	0.899	0.58
MSgM1-2	CL	0.643	0.70
MSgM1-3	CL	0.641	0.55
MSgM1-10	CL	1.161	0.70
MSgM2-23m	ML	0.930	0.20
Artificially reconstructed mixtures			
M1	-	0.816	0.01
M2	-	0.680	0.21
M3	-	0.851	0.27
M4	-	0.843	0.34
M5	-	1.022	0.50
M6	-	1.230	0.67
Reconstituted samples			
MSgM2-9b	ML	0.780-1.270	0.04
MSgM2-13b	ML	0.570-0.195	0.11
MSgM2-25b	ML	0.635-0.897	0.10
MSgM2-18a	SP-SM	0.686-0.980	0.01
MSgM2-19	SP-SM	0.725-0.981	0.01
MSgM2-21b	SP-SM	0.645-0.889	0.01
MSgM2-23a	SP-SM	0.731-1.022	0.01

3.3. PROBABILSTIC UNIFIED SOIL COMPRESSION MODEL

The probabilistic model is specifically developed for the unified soil compression model to eliminate the bias in the original deterministic estimates and properly account for the

uncertainties in the prediction process. The model parameters e_{g1v}, e_{c1v}, ρ_c, and α in Eq. (3.5) are replaced by a vector of unknown model parameters $\boldsymbol{\theta} = (e_{g1v}, e_{c1v}, \rho_c, \alpha)$ and an additive error term, $\sigma\varepsilon_i$, is considered to account for the uncertainty in the model. In the probabilistic model for each compression curve the mean void ratio μ_{e_i} at a given vertical stress p'_{vi} depends on the mean void ratio and stress computed at the previous loading step $(\mu_{e_{i-1}}, p'_{vi-1})$ according to the following form:

$$e_i(\boldsymbol{\theta},\sigma) = \mu_{e_{i-1}}\left\{1-\rho_c\left\{1-\left[1-\frac{p'_{vi}}{p_{at}}\left(\frac{\mu_{e_{i-1}}}{e_{1v}\left(e_{g1v},e_{c1v},FF\right)}\right)^{1/\rho_c}\right]^{\alpha}\right\}\frac{\left(p'_{vi-1}-p'_{vi}\right)}{p'_{vi}}\right\}+\sigma\varepsilon_i,\quad i=1,\ldots,n_k \tag{3.6}$$

where $(\boldsymbol{\theta},\sigma)$ = a set of unknown model parameters, ε_i = random variable with zero mean and unit variance, σ = the standard deviation of the model error. An additional model parameter is the coefficient of correlation ρ between two values of the errors ε_i and ε_j within the same test. Here we assume that data points on two different compressions curves are uncorrelated. Therefore, the covariance matrix for the variables ε_i in m different tests each with n_k data points per test is written as

$$\boldsymbol{\Sigma} = \begin{bmatrix} \boldsymbol{\Sigma}_1 & & & & \text{sym.} \\ \vdots & \ddots & & & \\ 0 & & \boldsymbol{\Sigma}_k & & \\ \vdots & & \vdots & \ddots & \\ 0 & \cdots & 0 & \cdots & \boldsymbol{\Sigma}_m \end{bmatrix}_{N\times N} \tag{3.7}$$

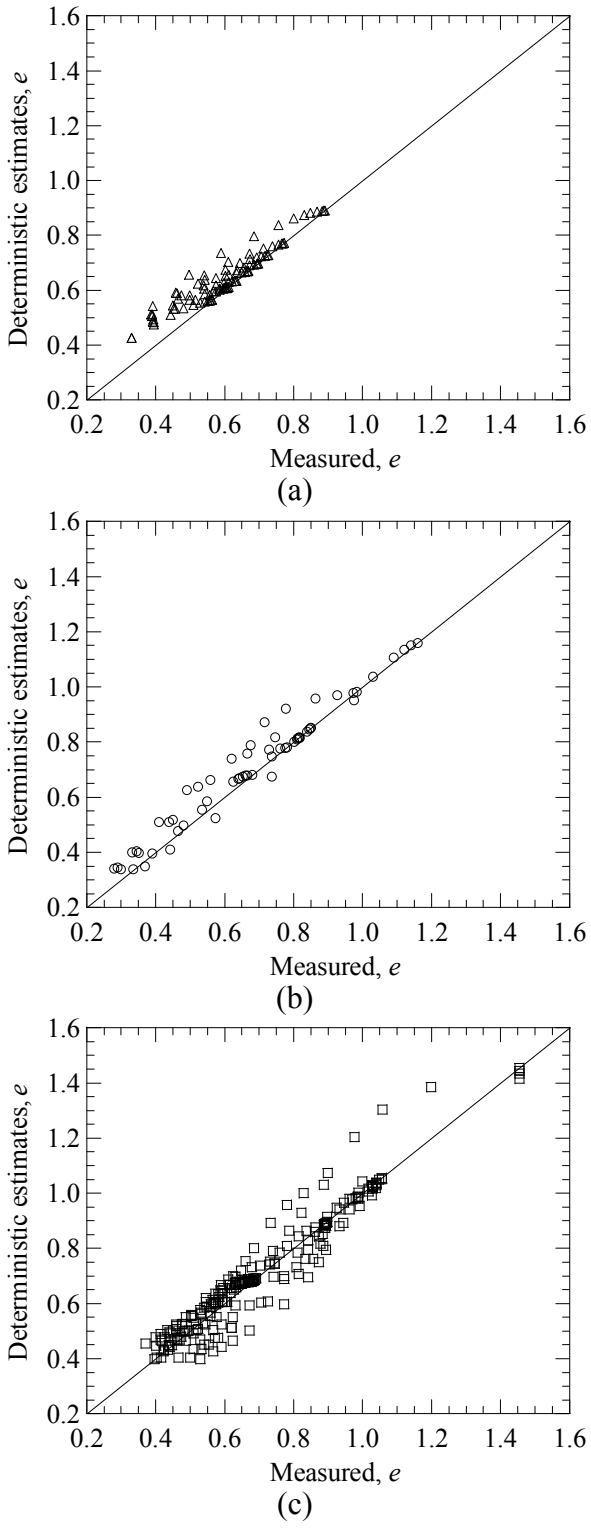


Fig. 3.2. Deterministic estimations with (a) the natural samples (b) the artificially reconstructed mixtures (c) the reconstituted samples

$$\text{where } \boldsymbol{\Sigma}_k = \begin{bmatrix} \sigma^2 & & & \text{sym.} \\ \rho\sigma^2 & \sigma^2 & & \\ \vdots & \vdots & \ddots & \\ \rho\sigma^2 & \rho\sigma^2 & \dots & \sigma^2 \end{bmatrix}_{n_k \times n_k}$$

where $N = \sum_{k=1}^{m} n_k$. Note that for given μ_e, p_v', p_{at}, $\boldsymbol{\theta}$, and σ, we have $\text{Var}[\Delta\mu_{e_i} / \mu_{e_{i-1}}] = \sigma^2$ as the variance of the model. In assessing the probabilistic model two assumptions are made: (1) the homoskedasticity assumption that assumes that σ is approximately constant over the range of the data used to assess the model, and (2) the normality assumption that assumes that ε is normally distributed. These assumptions can be verified using diagnostic plots of the data or the residuals against model predictions or individual regressors (Rao and Toutenburg, 1997). As a result, the unknown model parameters are only six $\boldsymbol{\Theta} = (e_{g1v}, e_{c1v}, \rho_c, \alpha, \sigma, \rho)$.

3.3.1. Bayesian Model Updating

In the proposed Bayesian approach, the unknown parameters $\boldsymbol{\Theta}$ are estimated by use of the updating rule (Box and Tiao 1992)

$$p\left(\boldsymbol{\Theta}|\mathbf{y}_1\right) \propto \kappa L\left(\boldsymbol{\Theta}|\mathbf{y}_1\right) p\left(\boldsymbol{\Theta}\right) \tag{3.8}$$

where $p(\boldsymbol{\Theta} \mid \mathbf{y}_1)$ is the posterior distribution representing our updated state of knowledge about $\boldsymbol{\Theta}$ based on the information in the observation $\mathbf{y}_1$, $L(\boldsymbol{\Theta} \mid \mathbf{y}_1)$ is the likelihood function representing the objective information on $\boldsymbol{\Theta}$ contained in the experimental data

$\mathbf{y}_1$, $p(\boldsymbol{\Theta})$ is the prior distribution reflecting our state of knowledge about $\boldsymbol{\Theta}$ prior to obtaining the experimental data, and $\kappa = [\int L(\boldsymbol{\Theta} \mid \mathbf{y}_1)\, p(\boldsymbol{\Theta}) d\boldsymbol{\Theta}\ ]^{-1}$ is a normalizing factor.

A Bayesian updating rule can be repeated to update a model any time new observations become available. For example, if a second set of observations $\mathbf{y}_2$, distributed independently of $\mathbf{y}_1$, are available, we can update $p(\boldsymbol{\Theta} \mid \mathbf{y}_1)$ with $\mathbf{y}_2$ using Eq. (3.9)

$$p\left(\boldsymbol{\Theta}|\mathbf{y}_1, \mathbf{y}_2\right) \propto L\left(\boldsymbol{\Theta}|\mathbf{y}_2\right) L\left(\boldsymbol{\Theta}|\mathbf{y}_1\right) p\left(\boldsymbol{\Theta}\right) \propto L\left(\boldsymbol{\Theta}|\mathbf{y}_2\right) p\left(\boldsymbol{\Theta}|\mathbf{y}_1\right) \tag{3.9}$$

where $p(\boldsymbol{\Theta} \mid \mathbf{y}_1)$ is now the prior distribution reflecting our state of knowledge before observing $\mathbf{y}_2$ and after observing $\mathbf{y}_1$.

Generalizing, given Q independent observations, we can write

$$p\left(\boldsymbol{\Theta}|\mathbf{y}_1, ..., \mathbf{y}_q\right) \propto L\left(\boldsymbol{\Theta}|\mathbf{y}_q\right) p\left(\boldsymbol{\Theta}|\mathbf{y}_1, ..., \mathbf{y}_{q-1}\right) \quad q=2,...,Q \tag{3.10}$$

where $p(\boldsymbol{\Theta} \mid \mathbf{y}_1)$ is given as in Eq. (3.8). Eq. (3.10) can be used to repeatedly update our current knowledge about the unknown parameters $\boldsymbol{\Theta}$, as new observations become available.

3.3.2. Prior Distribution

The prior distribution should be constructed using all available knowledge, including previous statistical analysis, and engineering judgment and expertise. If there is no existing information related to the unknown parameters $\boldsymbol{\Theta}$, a noninformative prior should be used reflecting that nothing or little is known a priori. In the case of a linear

model in the unknown parameters $\boldsymbol{\Theta}$, a uniform prior can be used as the noninformative prior (Box and Tiao, 1992). However, when a probabilistic model is a nonlinear function of the unknown parameters $\boldsymbol{\Theta}$, a uniform distribution might not be noninformative. In this case, an approximate noninformative prior can be developed using Jeffreys' rule (1961). The assumption of independence between $\boldsymbol{\theta}$ and $\boldsymbol{\Sigma}$ is made, i.e.,

$$p(\boldsymbol{\Theta}) \propto p(\boldsymbol{\theta})\, p(\boldsymbol{\Sigma}) \tag{3.11}$$

According to Jeffreys' rule, the prior distribution of the unknown parameters $\boldsymbol{\theta}$ is proportional to the positive square root of the determinant of the information matrix. The information matrix is the expectation of the negative second partial derivative of the log–likelihood function with respected to the unknown parameters $\boldsymbol{\theta}$. Hence, the Jeffreys' prior for $\boldsymbol{\theta}$ can be written as

$$p(\boldsymbol{\theta}) \propto \left| \mathscr{I}_n(\boldsymbol{\theta})_{ij} \right|^{-1/2} = \left| E\left[-\frac{\partial^2 \log L}{\partial \boldsymbol{\theta}_i \partial \boldsymbol{\theta}_j} \right] \right|^{-1/2} \tag{3.12}$$

where $|\cdot|$ = the determinant. Furthermore, Gardoni et al. (2002a) have shown that the noninformative prior for (σ, ρ) can be written as

$$p(\sigma, \rho) \propto |\mathbf{R}|^{-(N+1)/2} \prod_{i=1}^{N} \frac{1}{\sigma_i} \tag{3.13}$$

where $\mathbf{R}$ = correlation matrix.

3.3.3. Likelihood Function and Maximum Likelihood Estimates

The likelihood function is proportional to the conditional probability of observing the laboratory measurements for given values of the model parameters. Based on the correlation structure in Eq. (3.7), the multivariate normal likelihood model can be written as

$$L\left(\boldsymbol{\Theta}\middle|\mathbf{e}\right)=L\left(\boldsymbol{\theta},\boldsymbol{\Sigma}\middle|\mathbf{e}\right)=\prod_{k=1}^{m}\left(2\pi\right)^{-n_k/2}\left|\boldsymbol{\Sigma}_k\right|^{-1/2}\exp\left[-\frac{1}{2}\boldsymbol{\varepsilon}_k^T\left(\boldsymbol{\theta}\right)\boldsymbol{\Sigma}_k^{-1}\boldsymbol{\varepsilon}_k\left(\boldsymbol{\theta}\right)\right] \tag{3.14}$$

where $\mathbf{e}$ = vector of observed void ratios e_i, and $\boldsymbol{\varepsilon}_k$ = vector of errors for each compression curve. To verify the normality assumption made for the model errors, Fig. 3.3 shows the Quantile-Quantile (Q-Q) plots (Chambers et al. 1983) for the three samples considered. It can be seen that the residuals closely follow a normal distribution. Before accounting for the prior information, the unknown model parameters $\boldsymbol{\Theta}$ are estimated based only on the laboratory measurements by maximizing the likelihood function. Table 3.2 lists the maximum likelihood estimates (MLE) of the parameters $\boldsymbol{\Theta}$ for the natural samples, the artificially reconstructed mixtures, and the reconstituted samples. In the MLE analysis for the reconstituted samples, there are only three unknown material parameters e_{c1v}, ρ_c, and α because the laboratory tests indicate that clays fractions FF are less than 0.2. The values of the parameters are maximized with the likelihood function in Eq. (3.14). Fig. 3.4 shows MLE estimates versus the laboratory measurements of the void ratio. For a perfect model, the data should line up along the 1:1 solid line. The dotted lines delimit the region within 1 standard deviation of the mean. The following observations can be made from these results:

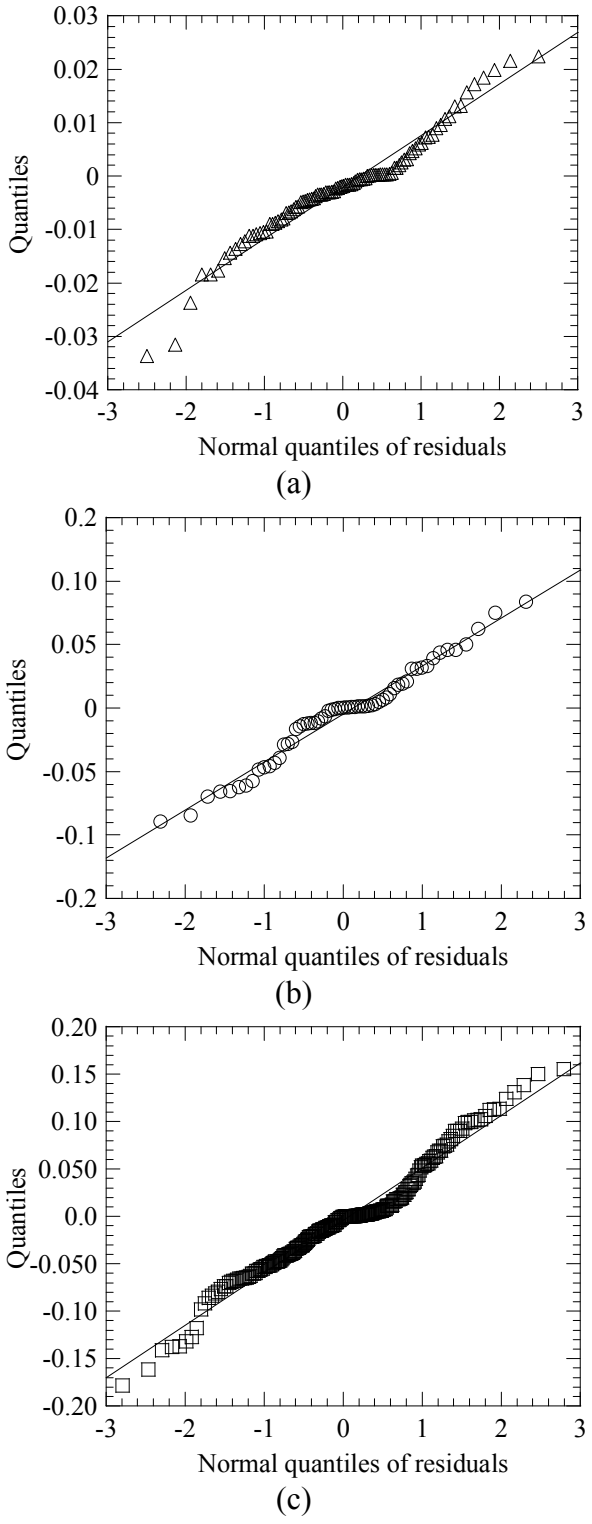


Fig. 3.3. Q-Q plots of the residuals for MLE estimations with (a) the natural samples (b) the artificially reconstructed mixtures (c) the reconstituted samples

Table 3.2. MLE of the unknown parameters

Type of sample	Parameter	Mean	Standard deviation	Correlation coefficient					
				e_{g1v}	e_{c1v}	ρ_c	α	σ	ρ
Natural samples	e_{g1v}	2.18	0.796	1					
	e_{c1v}	2.22	0.792	0.99	1				
	ρ_c	0.297	0.087	0.94	0.93	1			
	α	1.61	1.03	–0.13	–0.09	–0.46	1		
	σ	0.009	0.001	–0.03	–0.05	0.12	–0.44	1	
	ρ	0.126	0.102	0.78	0.76	0.88	–0.54	0.17	1
Artificially reconstructed mixtures	e_{g1v}	1.44	0.009	1					
	e_{c1v}	1.77	0.103	0.78	1				
	ρ_c	0.199	0.013	0.92	0.83	1			
	α	0.935	0.259	0.74	0.72	0.63	1		
	σ	0.033	0.003	–0.01	0.01	–0.02	0.04	1	
	ρ	0.194	0.065	0.11	0.09	0.08	0.04	0.00	1
Reconstituted samples	e_{g1v}	2.47	2.23	1					
	e_{c1v}	-	-	-	-				
	ρ_c	0.221	0.33	0.99	-	1			
	α	5.24	1.06	0.05	-	0.03	1		
	σ	0.044	0.122	–0.98	-	–0.98	–0.03	1	
	ρ	0.195	0.031	–0.15	-	–0.15	–0.00	0.14	1

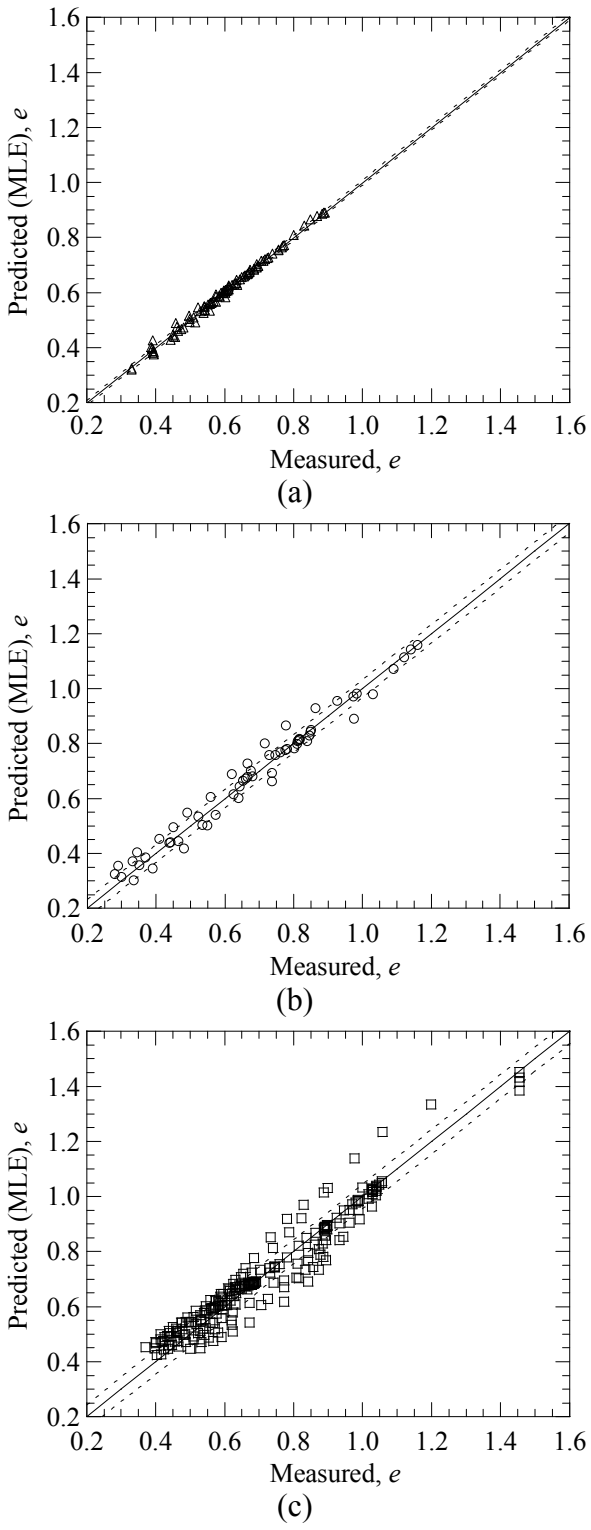


Fig. 3.4. MLE with (a) the natural samples (b) the artificially reconstructed mixtures (c) the reconstituted samples

1. The standard deviation associated with the model prediction is 0.009 for the natural samples, 0.033 for the artificially reconstructed mixtures, and 0.044 for the reconstituted samples. These indicate that there is more uncertainty with the database in the artificially reconstructed mixtures and the reconstituted samples than in the natural samples. This is in part due to the larger size of the reconstituted sample.
2. As mentioned before, the parameters e_{g1v} and e_{c1v} are determined from FF. The standard deviation of e_{g1v} and e_{c1v} for the artificially reconstructed mixtures is smaller than the one based on the natural samples due to the fact that the FF ranges from 0.01 to 0.67 in the mixtures while the FF is in the range of 0.2–0.78 in the natural samples (Table 3.1).
3. The probabilistic compression model improves the model estimates compared to the deterministic model and removes the inherent bias.

3.3.4. Posterior Estimates

According to the general Bayesian approach, posterior estimates represent our updated state of knowledge about the compression model. In this section, posterior statistics of $\boldsymbol{\Theta}$ are first computed assuming a noninformative prior and using the natural samples. The posterior statistics are then updated first using the artificially reconstructed mixtures, and then further updated using the reconstituted samples. Table 3.3 lists the posterior statistics. Fig. 3.5 shows the posterior estimates versus the laboratory measurements based on mean values of the unknown parameters for the natural samples.

Table 3.3. Posterior statistics of the unknown parameters for the natural samples

Parameter	Mean	Standard deviation	Correlation coefficient					
			e_{g1v}	e_{c1v}	ρ_c	α	σ	ρ
e_{g1v}	2.12	0.284	1					
e_{c1v}	2.16	0.298	0.99	1				
ρ_c	0.295	0.015	0.94	0.94	1			
α	1.51	0.489	0.94	0.95	0.78	1		
σ	0.009	0.001	–0.04	–0.04	–0.03	–0.04	1	
ρ	0.126	0.047	–0.06	–0.07	–0.01	–0.13	0.01	1

The posterior means of the parameters are similar to the MLE ones based on the same samples.

Starting from the posterior estimates based on the natural samples, we use Eq. (3.9) to update the posterior with data from the reconstructed samples. Table 3.4 lists the updated posterior statistics and Fig. 3.6 shows the posterior predictions based on both the natural and reconstituted samples for the probabilistic compression model. The

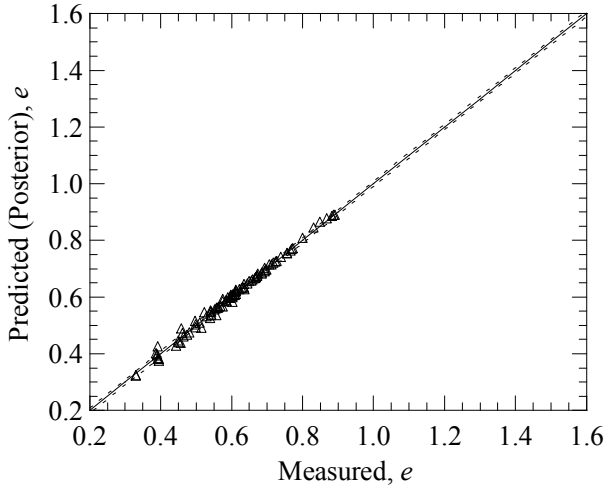


Fig. 3.5. Posterior estimates with the natural samples

Table 3.4. Updated posterior statistics of the unknown parameters for the natural samples and the artificially reconstructed mixtures

Parameter	Mean	Standard deviation	Correlation coefficient					
			e_{g1v}	e_{c1v}	ρ_c	α	σ	ρ
e_{g1v}	2.09	0.332	1					
e_{c1v}	2.15	0.337	0.98	1				
ρ_c	0.251	0.011	0.73	0.71	1			
α	2.73	1.58	0.94	0.95	0.48	1		
σ	0.027	0.002	–0.23	–0.23	–0.11	–0.25	1	
ρ	0.195	0.039	0.06	0.04	–0.03	0.08	–0.03	1

updated posterior estimates reflect both the information content of the old and of the new data. The values of the updated posterior means of the unknown parameters are between the posterior estimates for the natural samples and the MLE for the artificially reconstructed mixtures. However, the values of unknown parameters $\boldsymbol{\theta}$ are closer to the posterior statistics of the natural samples (100 data points) than the MLE for the

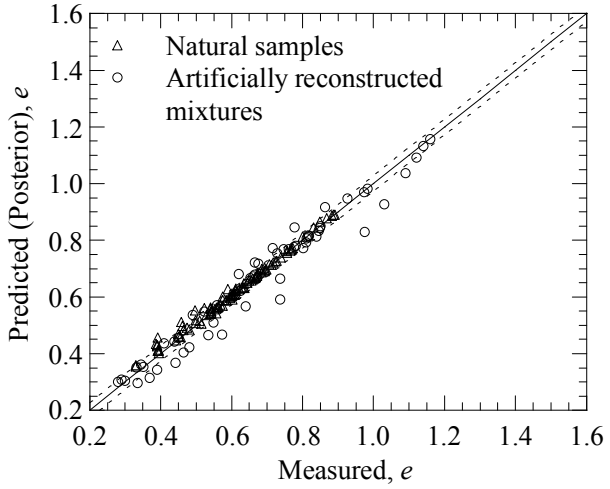


Fig. 3.6. Updated posterior estimates with the natural samples and the artificially reconstructed mixtures

Table 3.5. Updated posterior statistics of the unknown parameters for three samples

Parameter	Mean	Standard deviation	Correlation coefficient					
			e_{g1v}	e_{c1v}	ρ_c	α	σ	ρ
e_{g1v}	2.39	1.76	1					
e_{c1v}	2.01	1.35	0.99	1				
ρ_c	0.226	0.022	0.96	0.96	1			
α	4.45	2.76	0.99	0.99	0.95	1		
σ	0.041	0.002	0.49	0.48	0.46	0.49	1	
ρ	0.199	0.032	–0.67	–0.67	–0.64	–0.67	–0.33	1

artificially reconstructed samples (60 data points). This is due to the larger sample size of the natural data. Due to the addition of new data with different information content, the updated model standard deviation increases.

An additional posterior estimation is also carried out accounting for the information from the reconstituted samples. Table 3.5 lists the posterior statistics associated with the three samples. The posterior standard deviations of the model

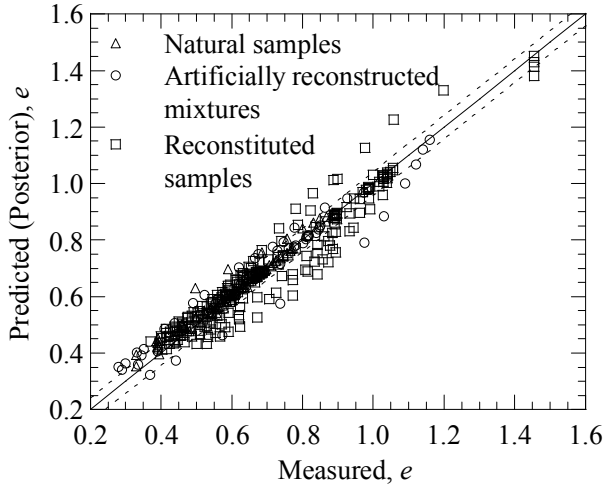


Fig. 3.7. Updated posterior estimates with three samples

parameters reflect the statistical uncertainty in the estimates. This uncertainty is epistemic and can be reduced by collecting more data. The values of the coefficient of variation (COV) suggest that there is considerable uncertainty about the values of α and, to a lesser extent, of e_{g1v}, e_{c1v}, ρ_c, and ρ. Fig. 3.7 shows the mean predicted versus the measured e for all the data. As expected, the values of the unknown parameters fall between posterior estimation for the natural samples and the MLE for the reconstituted samples but closer to the MLE based on the reconstituted samples (238 data points).

The results of the probabilistic compression model show that the model parameter ρ_c is overestimated in the deterministic model and the parameters e_{g1v}, e_{c1v}, and α are underestimated.

Table 3.6. *MAPE* values for all the models

Model	Type of sample	*MAPE* (%)
Deterministic estimations	Natural samples	8.865
	Artificially reconstructed mixtures	7.246
	Reconstituted samples	6.354
MLE	Natural samples	1.332
	Artificially reconstructed mixtures	5.071
	Reconstituted samples	6.224
Posterior estimates	Natural samples	1.332
	Natural samples and artificially reconstructed mixtures	3.009
	Natural samples, artificially reconstructed mixtures, and reconstituted samples	5.209

The accuracy of the models can be assessed using the Mean Absolute Percent Error (*MAPE*) defined as

$$MAPE = \frac{1}{n}\left[\sum_{i=1}^{n}\frac{\left|\hat{e}_i - e_i\right|}{e_i}\right]\times 100 \tag{3.15}$$

where $\hat{e}_i$ = the predicted value of e_i for each data point, and n = the number of observations used to assess the models. The *MAPE* indicates the average relative error and is a measure of the accuracy of model predictions. Table 6 lists the values of the *MAPE* computed using the deterministic, MLE, and posterior estimates. It can be seen that the values of *MAPE* for the probabilistic models are lower than those of the deterministic model indicating that the proposed probabilistic models are more accurate.

3.4. CONCLUSIONS

Both the compression response and its variability must be characterized in order to estimate the probability of a differential settlement exceeding a certain threshold value. The work presented in this research introduces a probabilistic framework to address this issue in a rigorous manner, while preserving the format of a typical geotechnical settlement analysis. In order to avoid dealing with different approaches for each category of soil, a unified compression model developed by Biscontin et al. (2007) was used to characterize the nonlinear compression behavior of soils of varying gradation through a single framework. The simplified version of the deterministic model adopted for this work relies on a small number of parameters for each soil with similar mineralogy, irrespective of initial water content, depth, or stress history. Although this

reduces the complexity in the probabilistic framework, the same approach can be adapted to deal with a more traditional consolidation geotechnical analysis. As shown in the chapter, the results from a traditional laboratory testing program are used to compute the statistics of the model parameters and the uncertainty inherent in the model. A Bayesian approach also provides a rigorous procedure for incorporating additional information and improving settlement estimates as more samples are tested.

The next step in the process of estimating differential settlement requires the evaluation of the geometry and initial conditions in the field. The probabilistic unified soil compression model developed in the chapter can then be used for an accurate reliability analysis that properly account for the prevailing uncertainties.

CHAPTER IV

RELIABILITY ANALYSIS OF SETTLEMENT

4.1. INTRODUCTION

In general, soil deposit to support foundation results settlements. Settlement is a serious problem for many structures and can cause significant damage. Current practice in deterministic analysis of settlement assumes that the soil profile at a site is uniform from location to location, and only allows limited considerations on the variations of material properties and initial conditions within soil layers in spite of the wide range of compositions, gradations, and water contents in natural soils. Traditionally, foundation settlements have been estimated using linear elastic soil models, or by Terzaghi's consolidation theory, and then calculating the settlement with a deterministic analytical model. These formulations are deterministic and tend to be conservative. More recently, Baecher and Ingra (1981), Zeitoun and Baker (1992), Brzakala and Pula (1996), Nour et al. (2002), and Fenton and Griffiths (2002 and 2005) proposed probabilistic models for settlement analysis . Existing analyses are based on the elastic theory whereas only few approaches considered the consolidation method. In addition, little research work has been done on the effect of reliability tools, variable correlation function and changing stress state (Elkateb et al., 2003). Finally, Jung et al. (2009) developed a probabilistic model to develop a framework able to account for uncertainties in a formal and logical manner. The model was constructed based on a deterministic model form developed by Biscontin et al. (2007), which reflects the engineering

understanding of the phenomenon. However this probabilistic model could not account for the variability of loading step in one-dimensional laboratory tests.

This paper builds a new probabilistic compression model with based on the model developed by Jung et al. (2009) and to account the limitation stated above. In order to account the limitation, the probabilistic compression model is formulated by using auto-regressive model. Then, a Bayesian methodology is used to develop an unbiased probabilistic model that accurately predicts the settlements and account for all the prevailing uncertainties.

As an application, the set of model parameters is calibrated using data from a well-documented case history in Venice, Italy. The database includes soil properties and the field settlements obtained from a full-scale test embankment constructed near Venice. A Bayesian approach is used for the model assessment accounting for both a prior knowledge based on the first set of data and additional data that became available at a later time. The developed model is unbiased and accounts for the inherent uncertainties.

The developed probabilistic model is used to assess the conditional probability (fragility) of exceeding a specified settlement threshold for a given vertical pressure. Predictive fragilities are developed with special attention given to the treatment and quantification of aleatory and epistemic uncertainties. Sensitivity and importance measures are carried out to identify to which parameter(s) and random variable(s) are key for the reliability of the settlements. The developed estimates provide a sound basis

for decision about the need to design for settlements and the type of intervention most suitable.

This paper is composed of four sections. Following this introduction, we develop the probabilistic model for soil compression using laboratory data and the Bayesian updating rule. Next, we estimate the settlement for the Treporti Test Site (TTS) in the Venice Lagoon. Then, the fragility estimates are developed. Finally, we assess the sensitivity and importance measures.

4.2. PROBABILITIC MODEL FOR SOIL COMPRESSION

This section introduces a brief review of the probabilistic compression model based on a unified compression model. Then we present the proposed probabilistic model by using the Bayesian updating rule with new formulation and additional laboratory data to estimate the differential settlement.

4.2.1. Available Probabilistic Compression Model

Biscontin et al. (2007) developed the unified compression model to describe the one-dimensional compression response. The model converges to a linear limiting compression curve (K_0-LCC) in a double logarithm of the void ratio and vertical effective stress ($\log e - \log p'_v$) at high stress level. The one-dimensional compression response of this model can be characterized using a few model parameters (ρ_c, e_{c1v}, e_{g1v}, and α). Among these model parameters, the K_0-LCC can be described by the

slope, ρ_c, which is the same irrespective of grain size for particular soil mineralogy and a reference void ratio, e_{1v}, at the atmospheric pressure, p_{at}. The parameter α controls the curvature of the compression curve in the transitional regime preceding the K_0-LCC. Also, the position of the K_0-LCC defined by e_{1v} is a function of the combination between clay fraction, FF, and the reference void ratio of the clay-water and granular phase, e_{c1v} and e_{g1v}, respectively. Where e_{g1v} = reference void ratio for the K_0-LCC of the granular phase only, and e_{c1v} = reference void ratio for the K_0-LCC of the clay-water phase only. As a result, this model is uniquely characterized by the mineralogy of the soils, and is also able to describe a wide range of responses.

Then, Jung et al. (2009) developed a probabilistic compression model including the parameters. The aim of this probabilistic model is (a) to eliminate the bias after correcting the input parameters in the original unified compression model, and (b) to properly account for the aleatory and epistemic uncertainties, in the prediction process. To account for the uncertainty in the model, additive error term, $\sigma\varepsilon_i$, was considered, where ε_i = random variable with zero mean and unit variance and σ = the standard deviation of the model error. In the model, the void ratio e_i at a given vertical stress, p'_{vi}, depends on the mean void ratio and stress computed at the previous loading step $(\mathrm{E}(e_{i-1}), p'_{vi-1})$ in the probabilistic model for each compression curve. According to Jung et al. (2009), the probabilistic compression model is written as

$$e_i(\boldsymbol{\theta},\sigma)=\mathrm{E}\left[e_{i-1}(\boldsymbol{\theta},\sigma)\right]\left\{1-\rho_c\left\{1-\left[1-\frac{p'_{vi}}{p_{\mathrm{at}}}\left(\frac{\mathrm{E}\left[e_{i-1}(\boldsymbol{\theta},\sigma)\right]}{e_{1v}\left(e_{g1v},e_{c1v},FF\right)}\right)^{1/\rho_c}\right]^{\alpha}\right\}\frac{(p'_{vi-1}-p'_{vi})}{p'_{vi}}\right\}+\sigma\varepsilon_i,\quad i=1,\ldots,n_k \quad (4.1)$$

where $\boldsymbol{\theta}=(e_{g1v},e_{c1v},\rho_{\mathrm{c}},\alpha)=$ a vector of unknown model parameters, $n_k=$ the number of data points in each one-dimensional compression curve, where *k* indicates a specific sample. In assessing the probabilistic model two assumptions are made: (1) the homoskedasticity assumption that assumes that σ is approximately constant over the range of the data used to assess the model, and (2) the normality assumption that assumes that ε is normally distributed.

Furthermore, the coefficient of correlation, ρ, between two values of the errors ε_i and ε_j within the same test was introduced and it was assumed that the predictions on two different compression curves are uncorrelated. Finally, the unknown model parameters were $\boldsymbol{\Theta}=(\boldsymbol{\theta},\sigma,\rho)$. A more detailed description of the probabilistic compression model can be found in Jung et al. (2009).

4.2.2. Formulation of the Proposed Probabilistic Compression Model

We now propose a new model to update the probabilistic compression accounting on the step of the loading in one-dimensional compression tests. In proposed model, we employ lognormal distribution of the void ratio and unit-median error term. Therefore, the updated probabilistic compression model can be written as

$$\ln\left[e_i(\boldsymbol{\theta},\sigma)\right]=\ln\left\{\mathrm{E}\left[e_i(\boldsymbol{\theta},\sigma)\right]\left\{1-\rho_c\left\{1-\left[1-\frac{p'_{vi}}{p_{at}}\left(\frac{\mathrm{E}\left[e_{i-1}(\boldsymbol{\theta},\sigma)\right]}{e_{1v}\left(e_{glv},e_{clv},FF\right)}\right)^{1/\rho_c}\right]^{\alpha}\left\{\frac{(p'_{vi-1}-p'_{vi})}{p'_{vi}}\right\}\right\}\right\}+\sigma_i\varepsilon_i,\quad i=1,\ldots,n_k \tag{4.2}$$

where $\sigma_i=\sigma(p_{vi}-p_{vi-1})^{\beta}=$ the standard deviation of the model error, $\boldsymbol{\Theta}=(\boldsymbol{\theta},\sigma,\beta,\rho)=$ a set of unknown model parameters, $\varepsilon_i=$ random variable with zero mean and unit variance. A logarithmic variance stabilizing transformation is used to satisfy the following two assumptions: (1) σ is constant (homoskedasticity assumption) and (2) ε_i are normally distributed with zero mean and unit standard deviation.

Then, we made the same assumption for the coefficient of correlation. Therefore, the covariance matrix for the variables ε_i in m different test each with n_k data points per test can be written

$$\boldsymbol{\Sigma}=\begin{bmatrix}\boldsymbol{\Sigma}_1 & & & \text{sym.}\\ 0 & \boldsymbol{\Sigma}_2 & & \\ \vdots & \vdots & \ddots & \\ 0 & 0 & \cdots & \boldsymbol{\Sigma}_k\end{bmatrix}_{N\times N} \tag{4.3}$$

$$\text{where } \boldsymbol{\Sigma}_k=\begin{bmatrix}\sigma_1^2 & & & \text{sym.}\\ \rho\sigma_2\sigma_1 & \sigma_2^2 & & \\ \vdots & \vdots & \ddots & \\ \rho\sigma_{n_k}\sigma_1 & \rho\sigma_{n_k}\sigma_2 & \cdots & \sigma_{n_k}^2\end{bmatrix}_{n_k\times n_k}$$

As a result, the unknown model parameters are only seven $\boldsymbol{\Theta}=(\boldsymbol{\theta},\sigma,\beta,\rho)=$ a set of unknown model parameters.

4.3. MODEL ASSESSMENT

The proposed approach to build the probabilistic compression model is based on the Bayesian updating rule. The proposed approach is capable of incorporating all types of available information, including mathematical models, laboratory data.

4.3.1. Bayesian Updating Rule

In the proposed Bayesian approach, the unknown parameters $\boldsymbol{\Theta}$ are estimated by use of the updating rule (Box and Tiao 1992)

$$p\left(\boldsymbol{\Theta}|\mathbf{y}\right) \propto \kappa L\left(\boldsymbol{\Theta}|\mathbf{y}\right) p\left(\boldsymbol{\Theta}\right) \tag{4.4}$$

where $p(\boldsymbol{\Theta}\,|\,\mathbf{y})$ is the posterior distribution representing our updated state of knowledge about $\boldsymbol{\Theta}$ based on the information in the observation $\mathbf{y}$, $L(\boldsymbol{\Theta}\,|\,\mathbf{y})$ is the likelihood function representing the objective information on $\boldsymbol{\Theta}$ contained in the experimental data $\mathbf{y}$, $p(\boldsymbol{\Theta})$ is the prior distribution that reflects our state of knowledge about $\boldsymbol{\Theta}$ prior to obtaining the experimental data, and $\kappa = [\int L(\boldsymbol{\Theta}\,|\,\mathbf{y})\, p(\boldsymbol{\Theta}) d\boldsymbol{\Theta}\,]^{-1}$ is a normalizing factor.

Due to the lack of prior information, we use a noniformative prior distribution constructed following Jung et al. (2009). Furthermore, details on the formulation of the likelihood function can also be found in Jung et al. (2009).

4.3.2. Laboratory Data

Three sets of natural samples, artificially reconstructed mixtures, and reconstituted samples on one-dimensional laboratory tests from Biscontin et al. (2007) were used for the assessment of probabilistic model in Jung et al. (2009). Among these data sets, Jung

Table 4.1. Ranges of the material properties from the data used in Jung et al. (2008) (A) and when including the new additional data (B)

Sample	Material property	Range (A)	Range (B)
Natural samples	Soil class	CL, ML	CL, ML, SP-SM
	e_0	0.643 – 1.161	0.622 – 1.161
	FF	0.2 – 0.78	0.01 – 0.78
Artificially reconstructed mixtures	Soil class	-	-
	e_0	0.680 – 1.230	0.680 – 1.230
	FF	0.67 – 0.01	0.67 – 0.01
Reconstituted samples	Soil class	ML, SP-SM	ML, SP-SM
	e_0	0.570 – 1.270	0.570 – 1.270
	FF	0.01 – 0.11	0.01 – 0.11

et al. selected 9 sets of the natural samples which reached the K_0-LCC regime and ρ_c, 6 sets of the artificially reconstructed mixtures, and 7 sets of the reconstituted samples. Therefore, the data was composed of 100 data points from the natural samples, 60 data points from artificially reconstructed mixtures, and 238 data points from the reconstituted samples.

We consider the 8 new data sets (84 points) from natural samples to update the probabilistic model. Table 4.1 shows a comparison between the ranges of the material properties of the soil used in Jung et al. (2009) and the new expanded ranges after including the new data sets.

4.3.3. Proposed Probabilistic Compression Model

The probabilistic compression model is updated using the new probabilistic model and the laboratory data. Table 4.2 lists the posterior statistics of the parameters $\boldsymbol{\Theta}$. We note that the standard deviations of the model parameters reflect the statistical uncertainty in the posterior estimates. This epistemic uncertainty can be reduced by collecting more data.

Fig. 4.1 shows the mean error predictions between measured and predicted void ratio using the results of the posterior statistics based on the proposed probabilistic compression model. The data are shown as dots (•). It is noted that for a perfect model, the data should line up along the zero error line. Also, the dotted lines indicate the region within $\pm\sigma$ in the proposed probabilistic model.

Table 4.2.Posterior statistics of the unknown parameters

Parameter	Mean	Standard deviation	Correlation coefficient						
			e_{g1v}	e_{c1v}	ρ_c	α	σ	β	ρ
e_{g1v}	2.63	0.239	1						
e_{c1v}	2.81	0.303	0.96	1					
ρ_c	0.233	0.021	0.66	0.57	1				
α	5.86	2.83	0.89	0.93	0.29	1			
σ	0.009	0.001	0.13	0.03	0.26	–0.01	1		
β	0.296	0.017	–0.11	–0.01	–0.31	0.05	–0.76	1	
ρ	0.709	0.053	0.06	0.04	–0.02	0.06	0.62	–0.02	1

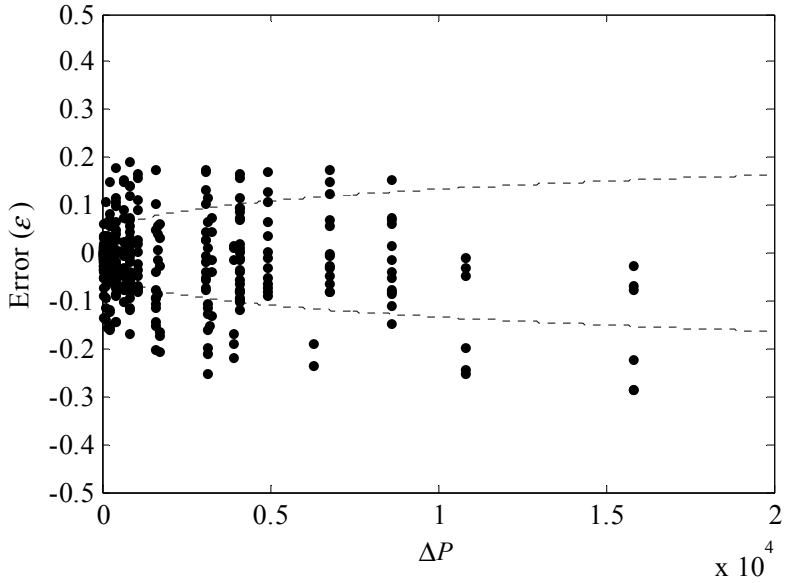


Fig. 4.1. Comparison between measured and predicted void ratio based on the proposed probabilistic model

4.4. SETTLEMENT ESTIMATES

In this section, we develop a probabilistic model to estimate the settlement of foundations using field/laboratory data and the proposed probabilistic compression model.

The estimates of the foundation settlements are constructed using the one-dimensional settlement. In general, the total settlement S can be calculated by summing the settlement in individual layers, ΔH_i, with varying compression properties or the initial void ratio with depth

$$S(\mathbf{x},\boldsymbol{\Theta}) = \sum_{i=1}^{r} \Delta H_i = \sum_{i=1}^{r} \left[\frac{\Delta e_i(\boldsymbol{\Theta})}{1+e_{0i}} H_{0i} \right] \tag{4.5}$$

where, $\Delta e_i(\boldsymbol{\Theta}) = e_{0i} - e_i(\boldsymbol{\Theta}) =$ the change in void ratio due to a given pressure P, $e_{0i} =$ initial void ratio, and $H_{0i} =$ the thickness of each soil layer, $\mathbf{x} = (P, e_{0i}, H_{0i}) =$ loading conditions, material properties and geometry that in general are random

As an application of the probabilistic model to estimate the foundation settlement, we employ the field and laboratory data presented by Simonini (2004). A number of successive and comprehensive geotechnical studies have been carried out to characterize the Venetian soils. Extensive work started at TTS with the field and laboratory characterization of the site by constructing a circular reinforced earth embankment reaching a 6.7 m height and 40 m diameter corresponding to about 106.5 kPa for a vertical pressure, P, at the end of construction. Simonini (2004) provided the site investigations, the laboratory investigations, and the monitoring of the vertical displacement under the center of the embankment. In the Eq. (4.5), the change in the void ratio can be calculated based on the probabilistic compression model using the initial conditions and the final conditions associated with the P. In order to compute the increase in stress, ΔP_v, due to the circular loading, Boussinesq's equation is used under the center of the embankment.

First, we determine the composition and the soil properties of the subsoil for the TTS. Then, the probabilistic compression model which aims to improve the accuracy of the settlement predictions by removing a potential bias associated with the prevailing uncertainties is used to estimate the foundation settlements. Table 4.3 lists the distribution model, mean, and standard deviation of $\mathbf{x}$. The mean values are based on the database (Simonini, 2004). The standard deviations are based on an assumed value

Table 4.3. Distribution model, mean, and standard deviation of a vector of measurable variables

Variable	Distribution	Mean	Standard deviation
P	Lognormal	106.5	10.65
e_{01}	Lognormal	1.10	0.11
e_{02}	Lognormal	1.10	0.11
e_{03}	Lognormal	1.05	0.11
e_{04}	Lognormal	1.00	0.10
e_{05}	Lognormal	1.27	0.13
e_{06}	Lognormal	1.05	0.11
e_{07}	Lognormal	0.89	0.09
e_{08}	Lognormal	0.89	0.09
e_{09}	Lognormal	0.91	0.09
e_{010}	Lognormal	0.87	0.09
e_{011}	Lognormal	1.02	0.10
e_{012}	Lognormal	0.87	0.09
e_{013}	Lognormal	0.90	0.09
e_{014}	Lognormal	1.14	0.11
e_{015}	Lognormal	0.90	0.09
e_{016}	Lognormal	1.01	0.10
e_{017}	Lognormal	0.95	0.10
e_{018}	Lognormal	0.90	0.09
e_{019}	Lognormal	0.91	0.09
e_{020}	Lognormal	0.96	0.10
e_{021}	Lognormal	0.94	0.09
e_{022}	Lognormal	0.89	0.09
e_{023}	Lognormal	0.82	0.08
e_{024}	Lognormal	0.78	0.08
e_{025}	Lognormal	0.84	0.08
e_{026}	Lognormal	0.91	0.09
e_{027}	Lognormal	1.51	0.15
e_{028}	Lognormal	0.92	0.09
e_{029}	Lognormal	1.10	0.11
e_{030}	Lognormal	1.10	0.11

Table 4.3. (*Continued.*)

Variable	Distribution	Mean	Standard deviation
e_{031}	Lognormal	0.88	0.09
e_{032}	Lognormal	1.02	0.10
e_{033}	Lognormal	0.95	0.10
e_{034}	Lognormal	0.93	0.09
e_{035}	Lognormal	0.88	0.09
e_{036}	Lognormal	0.98	0.10
e_{037}	Lognormal	1.22	0.12
e_{038}	Lognormal	1.15	0.12
e_{039}	Lognormal	1.14	0.11
e_{040}	Lognormal	1.03	0.10
e_{041}	Lognormal	0.94	0.09
e_{042}	Lognormal	1.08	0.11
e_{043}	Lognormal	1.04	0.10
e_{044}	Lognormal	1.20	0.12
e_{045}	Lognormal	0.86	0.09
e_{046}	Lognormal	0.98	0.10
e_{047}	Lognormal	1.05	0.11
e_{048}	Lognormal	1.08	0.11
e_{049}	Lognormal	1.33	0.13
e_{050}	Lognormal	1.24	0.12
e_{051}	Lognormal	1.32	0.13
e_{052}	Lognormal	0.83	0.08
e_{053}	Lognormal	0.85	0.09
e_{054}	Lognormal	0.90	0.09
e_{055}	Lognormal	1.04	0.10
e_{056}	Lognormal	0.85	0.09
e_{057}	Lognormal	1.03	0.10
e_{058}	Lognormal	0.83	0.08
e_{059}	Lognormal	0.99	0.10
e_{060}	Lognormal	1.02	0.10
e_{061}	Lognormal	0.91	0.09
e_{062}	Lognormal	1.00	0.10

Table 4.3. (*Continued.*)

Variable	Distribution	Mean	Standard deviation
e_{063}	Lognormal	0.98	0.10
e_{064}	Lognormal	0.90	0.09
e_{065}	Lognormal	1.01	0.10
e_{066}	Lognormal	0.94	0.09
e_{067}	Lognormal	1.08	0.11
e_{068}	Lognormal	1.02	0.10
e_{069}	Lognormal	1.13	0.11
e_{070}	Lognormal	0.84	0.08
e_{071}	Lognormal	0.77	0.08
e_{072}	Lognormal	0.84	0.08
e_{073}	Lognormal	0.99	0.10
e_{074}	Lognormal	0.92	0.09
e_{075}	Lognormal	0.89	0.09
e_{076}	Lognormal	0.98	0.10
e_{077}	Lognormal	1.00	0.10
e_{078}	Lognormal	0.75	0.08
e_{079}	Lognormal	0.88	0.09
e_{080}	Lognormal	0.79	0.08
e_{081}	Lognormal	0.90	0.09
e_{082}	Lognormal	1.03	0.10
e_{083}	Lognormal	1.21	0.12
e_{084}	Lognormal	0.78	0.08
e_{085}	Lognormal	0.86	0.09
e_{086}	Lognormal	0.91	0.09
e_{087}	Lognormal	1.03	0.10
e_{088}	Lognormal	0.96	0.10
e_{089}	Lognormal	0.92	0.09
e_{090}	Lognormal	0.90	0.09
e_{091}	Lognormal	1.40	0.14
e_{092}	Lognormal	1.14	0.11
e_{093}	Lognormal	1.10	0.11
e_{094}	Lognormal	1.17	0.12

Table 4.3. (*Continued.*)

Variable	Distribution	Mean	Standard deviation
e_{095}	Lognormal	1.19	0.12
e_{096}	Lognormal	0.86	0.09
e_{097}	Lognormal	0.95	0.10
e_{098}	Lognormal	0.95	0.10
e_{099}	Lognormal	0.89	0.09
H_{01}	Lognormal	0.50	0.05
H_{02}	Lognormal	0.63	0.06
H_{03}	Lognormal	0.53	0.05
H_{04}	Lognormal	0.30	0.03
H_{05}	Lognormal	0.25	0.03
H_{06}	Lognormal	0.25	0.03
H_{07}	Lognormal	0.65	0.07
H_{08}	Lognormal	1.05	0.11
H_{09}	Lognormal	0.80	0.08
H_{010}	Lognormal	0.75	0.08
H_{011}	Lognormal	0.80	0.08
H_{012}	Lognormal	0.50	0.05
H_{013}	Lognormal	0.50	0.05
H_{014}	Lognormal	0.45	0.05
H_{015}	Lognormal	0.25	0.03
H_{016}	Lognormal	0.25	0.03
H_{017}	Lognormal	0.70	0.07
H_{018}	Lognormal	1.20	0.12
H_{019}	Lognormal	1.05	0.11
H_{020}	Lognormal	0.80	0.08
H_{021}	Lognormal	1.20	0.12
H_{022}	Lognormal	1.40	0.14
H_{023}	Lognormal	0.90	0.09
H_{024}	Lognormal	0.85	0.09
H_{025}	Lognormal	1.20	0.12
H_{026}	Lognormal	0.75	0.08

Table 4.3. (*Continued.*)

Variable	Distribution	Mean	Standard deviation
H_{027}	Lognormal	0.25	0.03
H_{028}	Lognormal	0.40	0.04
H_{029}	Lognormal	0.50	0.05
H_{030}	Lognormal	0.45	0.05
H_{031}	Lognormal	0.35	0.04
H_{032}	Lognormal	0.30	0.03
H_{033}	Lognormal	0.35	0.04
H_{034}	Lognormal	0.50	0.05
H_{035}	Lognormal	0.55	0.06
H_{036}	Lognormal	0.35	0.04
H_{037}	Lognormal	0.20	0.02
H_{038}	Lognormal	0.20	0.02
H_{039}	Lognormal	0.30	0.03
H_{040}	Lognormal	0.40	0.04
H_{041}	Lognormal	0.40	0.04
H_{042}	Lognormal	0.40	0.04
H_{043}	Lognormal	0.50	0.05
H_{044}	Lognormal	0.60	0.06
H_{045}	Lognormal	0.50	0.05
H_{046}	Lognormal	0.30	0.03
H_{047}	Lognormal	0.20	0.02
H_{048}	Lognormal	0.30	0.03
H_{049}	Lognormal	0.35	0.04
H_{050}	Lognormal	0.25	0.03
H_{051}	Lognormal	0.25	0.03
H_{052}	Lognormal	0.35	0.04
H_{053}	Lognormal	0.30	0.03
H_{054}	Lognormal	0.30	0.03
H_{055}	Lognormal	0.50	0.05
H_{056}	Lognormal	0.55	0.06
H_{057}	Lognormal	0.45	0.05
H_{058}	Lognormal	0.30	0.03

Table 4.3. (*Continued.*)

Variable	Distribution	Mean	Standard deviation
H_{059}	Lognormal	0.25	0.03
H_{060}	Lognormal	0.25	0.03
H_{061}	Lognormal	0.30	0.03
H_{062}	Lognormal	0.55	0.06
H_{063}	Lognormal	0.65	0.07
H_{064}	Lognormal	0.40	0.04
H_{065}	Lognormal	0.20	0.02
H_{066}	Lognormal	0.20	0.02
H_{067}	Lognormal	0.30	0.03
H_{068}	Lognormal	0.70	0.07
H_{069}	Lognormal	0.60	0.06
H_{070}	Lognormal	0.35	0.04
H_{071}	Lognormal	0.45	0.05
H_{072}	Lognormal	0.40	0.04
H_{073}	Lognormal	0.40	0.04
H_{074}	Lognormal	0.30	0.03
H_{075}	Lognormal	0.30	0.03
H_{076}	Lognormal	0.40	0.04
H_{077}	Lognormal	0.40	0.04
H_{078}	Lognormal	0.45	0.05
H_{079}	Lognormal	0.45	0.05
H_{080}	Lognormal	0.45	0.05
H_{081}	Lognormal	0.55	0.06
H_{082}	Lognormal	0.40	0.04
H_{083}	Lognormal	0.20	0.02
H_{084}	Lognormal	0.20	0.02
H_{085}	Lognormal	0.20	0.02
H_{086}	Lognormal	0.35	0.04
H_{087}	Lognormal	0.50	0.05
H_{088}	Lognormal	0.60	0.06
H_{089}	Lognormal	0.70	0.07
H_{090}	Lognormal	0.65	0.07

Table 4.3. (*Continued.*)

Variable	Distribution	Mean	Standard deviation
H_{091}	Lognormal	0.60	0.06
H_{092}	Lognormal	0.50	0.05
H_{093}	Lognormal	0.45	0.05
H_{094}	Lognormal	0.65	0.07
H_{095}	Lognormal	0.65	0.07
H_{096}	Lognormal	0.65	0.07
H_{097}	Lognormal	1.10	0.11
H_{098}	Lognormal	1.30	0.13
H_{099}	Lognormal	1.50	0.15

for the coefficient of variation (COV) which reflects the confidence in the accuracy of the results of the site investigation. We use 10% COV.

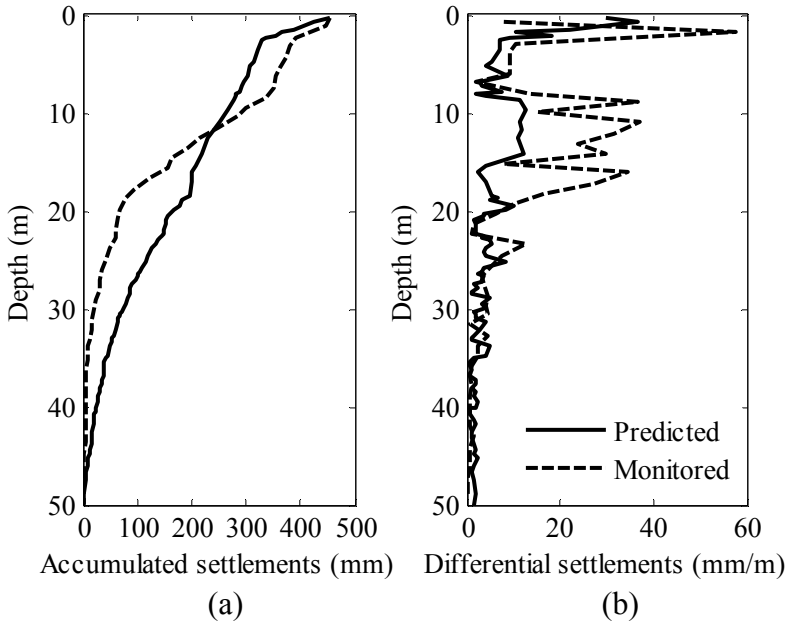


Fig. 4.2. Comparison between the predicted settlement and the monitored results for (a) accumulated and (b) differential settlements in TTS

On the basis of the probabilistic compression model and Eq. (4.5), Fig. 4.2 compares the predicted median settlement and the monitored results for accumulated and differential settlements in TTS. The computed overall settlement is closed to the monitored results. However, the estimates of the differential settlement in the depths 3 m and 10~20 m are lower than those monitored in the field. The reason for this might be that the probabilistic model is based on the Malamocco Test Sits (MTS) whereas the soil properties which show the same mineralogy characteristics with MTS are based on the TTS, approximately 20 km from TTS but still in Venice Lagoon.

4.5. FRAGILITY ASSESSMENT OF FOUNDATION SETTLEMENT

Fragility is defined as the conditional probability of attaining or exceeding prescribed limit states for a given set of demand variables. In this proposed work, we estimate the conditional probability (fragility) of attaining or exceeding a specified settlement threshold for a given value of the vertical pressure P. Following the conventional notation in structural reliability theory (Ditlevsen and Madsen, 1996), let $g(\mathbf{x},\boldsymbol{\Theta})$ be a mathematical model describing the limit state of interest. The limit state function, $g(\mathbf{x},\boldsymbol{\Theta})$, is defined such that the event $\{g(\mathbf{x},\boldsymbol{\Theta}) \le 0\}$ denotes the attainment or exceedance of the limit state. Where $\mathbf{x}$ denotes a vector of measurable variables and $\boldsymbol{\Theta}$ denotes a vector of model parameters.

Using the settlement model in Eq. (4.5) and considering a specified threshold settlement, S_T, the limit state functions can be formulated as

$$g\left(\mathbf{x},\boldsymbol{\Theta}\right) = S_T - S\left(\mathbf{x},\boldsymbol{\Theta}\right) \tag{4.6}$$

The fragility of a foundation can be formulated as

$$F\left(S_T,P,\boldsymbol{\Theta}\right) = \mathrm{P}\left[\left\{g\left(\mathbf{x},\boldsymbol{\Theta}\right) \le 0\right\} \middle| S_T,P\right] \tag{4.7}$$

where $\mathrm{P}[A\,|\,\mathbf{b}]$ denotes the conditional probability of event A for the given values of variable(s) $\mathbf{b}$. The uncertainty in the event for the given S_T and P arises from the inherent randomness in the variables $\mathbf{x}$, the inexact nature of the settlement model $S(\mathbf{x},\boldsymbol{\Theta})$ (or its sub-models), and the uncertainty inherent in the model parameters $\boldsymbol{\Theta}$.

To incorporate the epistemic uncertainties inherent in the model parameters $\boldsymbol{\Theta}$, we consider $\boldsymbol{\Theta}$ as random variables. The predictive estimate of fragility, $\tilde{F}(S_T,P)$, is

the expected value of $F(S_T, P, \boldsymbol{\Theta})$ over the posterior distribution of $\boldsymbol{\Theta}$ (Gardoni et al. 2002b), i.e.,

$$\tilde{F}\left(S_T, P\right) = \int F\left(S_T, P, \boldsymbol{\Theta}\right) f(\boldsymbol{\Theta})\, d\boldsymbol{\Theta} \tag{4.8}$$

where $f(\boldsymbol{\Theta})$= the posterior probability density function of $\boldsymbol{\Theta}$.

The Mote Carlo simulation (MC) and first order reliability method (FORM) is used in this study to estimate the fragility $\tilde{F}(S_T, P)$, since a closed-form solution of Eq. (4.8) is generally not available.

The probability of exceedance for the settlement applied to TTS is a function of the vertical pressure, P, and of the specified threshold settlement, S_T. Fig. 4.3 represents a conceptual three dimensional plot that show the probability of exceedance versus P and S_T. This figure shows that for a given P the probability of exceedance

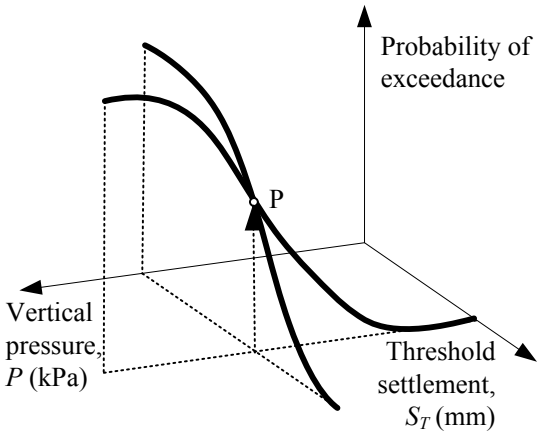


Fig. 4.3. Conceptual plot of the probability of exceedance, the vertical pressure and the specified threshold settlement

decrease as S_T increases, and for a given S_T the probability of exceedance increase as P increase.

4.5.1. Bounds on fragility

It is desirable to determine the contribution to the uncertainty in the fragility using form estimate due to the epistemic uncertainties. This uncertainty is reflected in the probability distribution of $F(S_T, P, \boldsymbol{\Theta})$ relative to the parameters $\boldsymbol{\Theta}$. Following Gardoni et al. (2002b), approximate confidence bounds can be written as

$$\left\{ \Phi\left[-\tilde{\beta}\left(S_T, P\right) - \sigma_{\beta}\left(S_T, P\right)\right],\ \Phi\left[-\tilde{\beta}\left(S_T, P\right) + \sigma_{\beta}\left(S_T, P\right)\right] \right\} \tag{4.9}$$

where $\Phi[\cdot] =$ the cumulative probability function of a standard normal random variable, $\beta(S_T, P, \boldsymbol{\Theta}) = \Phi^{-1}[1 - F(S_T, P, \boldsymbol{\Theta})] =$ the generalized reliability index, and

$$\sigma_{\beta}^2\left(S_T, P\right) \approx \nabla_{\boldsymbol{\Theta}} \beta\left(S_T, P\right) \boldsymbol{\Sigma}_{\boldsymbol{\Theta}\boldsymbol{\Theta}} \nabla_{\boldsymbol{\Theta}} \beta\left(S_T, P\right)^{\mathrm{T}} \tag{4.10}$$

where $\nabla_{\boldsymbol{\Theta}} \beta(S_T, P) =$ the gradient row vector of $\beta(S_T, P, \boldsymbol{\Theta})$ at the mean point $\mathbf{M}_{\boldsymbol{\Theta}}$, and $\boldsymbol{\Sigma}_{\boldsymbol{\Theta}\boldsymbol{\Theta}} =$ the posterior covariance matrix of $\boldsymbol{\Theta}$.

Fig. 4.4 shows the probability of exceedance with respect to (a) S_T at $P = 106.5$ kPa and (b) P at $S_T = 450$ mm. The circles indicate the predictive fragilities using MC), the dashed lines indicate the predictive fragilities using FORM and the dotted lines represent the approximate confidence bounds. It can be observed that the probabilities of exceedance estimated by MC are very close to ones estimated by FORM since we

have only one limit state. The dispersion indicated by the confidence bounds represent the influence of the epistemic uncertainty in the model parameters $\boldsymbol{\Theta}$.

4.6. SENSITIVITY AND IMPORTANCE MEASURES

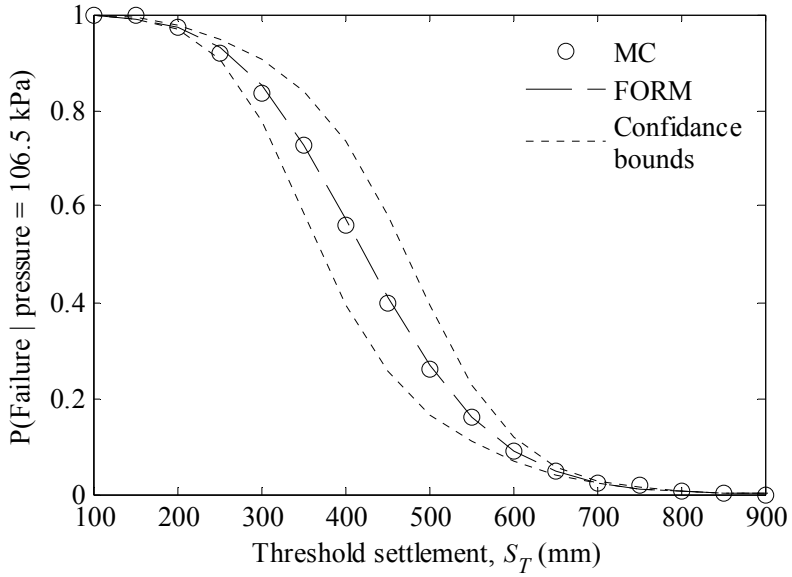
In this section, first we compute the sensitivity measures for the parameters used in the estimate of the foundation settlement. Then, we assess the importance measures for all random variable in the probabilistic model. We note that the sensitivity and importance measures are computed by using FORM.

4.6.1. Sensitivity Measures

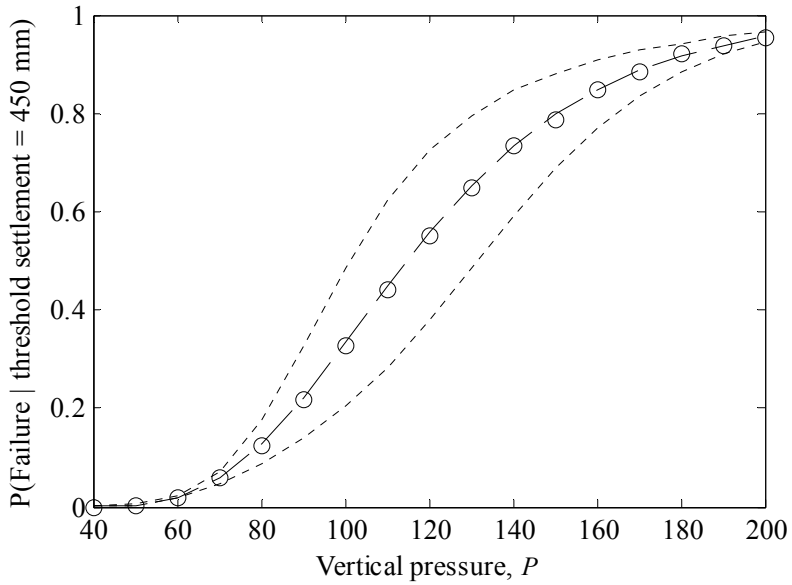
In the reliability assessment, sensitivity analysis is used to determine the effects on the reliability of the foundation settlement of changes in the parameters in the limit state function or in the distribution of the random variables. We can consider the influence on β and p of unit changes in $\mathrm{E}(\boldsymbol{\Theta})$ and $\mathrm{E}(\mathbf{x})$. where $\mathrm{E}(\cdot)$ indicates the mean of a random variable. The sensitivity measures can be computed following Hohenbichler and Rackwitz (1986). The gradient of the first-order reliability approximation of the failure probability is obtained by using the chain rule of the differentiation as

$$\nabla_{[\mathrm{E}(\boldsymbol{\Theta}),\mathrm{E}(\mathbf{x})]} p = -\varphi(\beta)\nabla_{[\mathrm{E}(\boldsymbol{\Theta}),\mathrm{E}(\mathbf{x})]}\beta \tag{4.11}$$

where $p =$ probability of the failure estimate using FORM, and $\varphi(\cdot) =$ the standard normal probability of density function, $\nabla_{[\mathrm{E}(\boldsymbol{\Theta}),\mathrm{E}(\mathbf{x})]}\beta =$ the derivative of the reliability index β with respect to $\mathrm{E}(\boldsymbol{\Theta})$ and $\mathrm{E}(\mathbf{x})$.



(a)



(b)

Fig. 4.4. Predictive fragility with confidence bounds as a function of (a) S_T at a given $P = 106.5$ kPa and (b) P at a given $S_T = 450$ mm

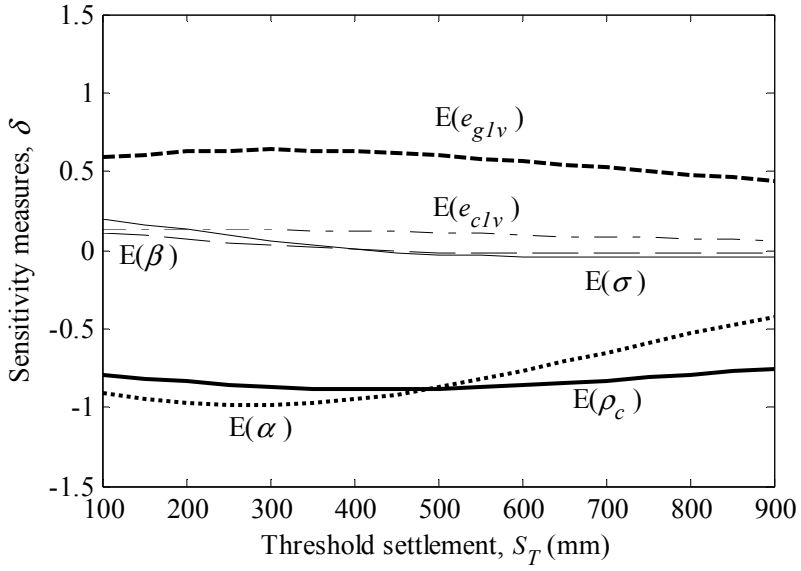
In this analysis, the means of the random variables and radius of the embankment, R, are considered as the parameters for the sensitivity analysis. Since we are interested in the sensitivities of β with respect to the mean of each random variable, we need to scale $\nabla_{\boldsymbol{\Theta}_g}\beta$ to compare the sensitivity measures of all parameters. On this basis we define the vector $\boldsymbol{\delta}$

$$\boldsymbol{\delta} = \mathbf{D}\nabla_{\boldsymbol{\Theta}_g}\beta \tag{4.12}$$

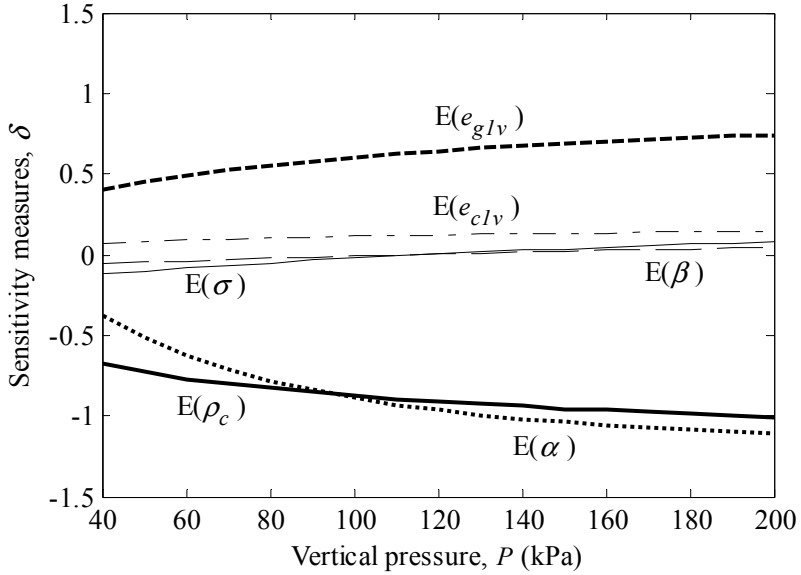
where $\mathbf{D} =$ the diagonal matrix with diagonal elements given by the standard deviation of each random variable. We note that the vector $\boldsymbol{\delta}$ renders the element of these vectors dimensionless and makes the parameter variations proportional to the corresponding standard deviations, which are measures of the underlying uncertainties..

Fig. 4.5(a) shows the sensitivity measures as a function of S_T for $P = 106.5$ kPa. $\mathrm{E}(\rho_c)$, $\mathrm{E}(\alpha)$, and $\mathrm{E}(e_{g1v})$ have larger effects on fragility while $\mathrm{E}(e_{c1v})$, $\mathrm{E}(\sigma)$, and $\mathrm{E}(\beta)$ have smaller effects We also computed the sensitivity measures for $\mathrm{E}(\mathbf{x})$ corresponding to each layer. The sensitivity measures with respect to them are negligible because they are basically zero. Therefore, we do not show them here. Among those parameters, $\mathrm{E}(\alpha)$ and $\mathrm{E}(\rho_c)$ are the most effective in increasing the specified threshold settlements.

Similarly, Fig. 4.5(b) shows the sensitivity measures as a function of P at a given $S_T = 450$ mm. The foundation settlement fragility is most sensitive to the means $\mathrm{E}(\rho_c)$ up to around 80 kPa of the vertical pressure. $\mathrm{E}(\alpha)$ then becomes the most sensitive parameter. In general, the positive sign of the sensitivity measure of each



(a)



(b)

Fig. 4.5. Sensitivity measures for the parameters as a function of (a) S_T at a given $P = 106.5$ kPa and (b) P at a given $S_T = 450$ mm

parameter indicate a “resistance” variable and the negative sign means a “load” variable in the limit state function.

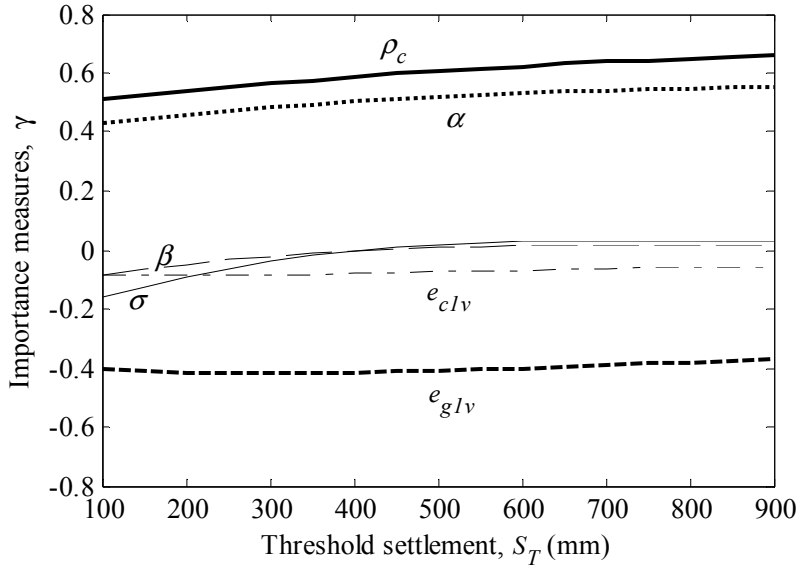
4.6.2. Importance Measures

We have several random variables which have different contributions to the variability of the limit state function defied in Eq. (4.6). Following Der Kiureghian and Ke (1995), the importance measure can be defined as

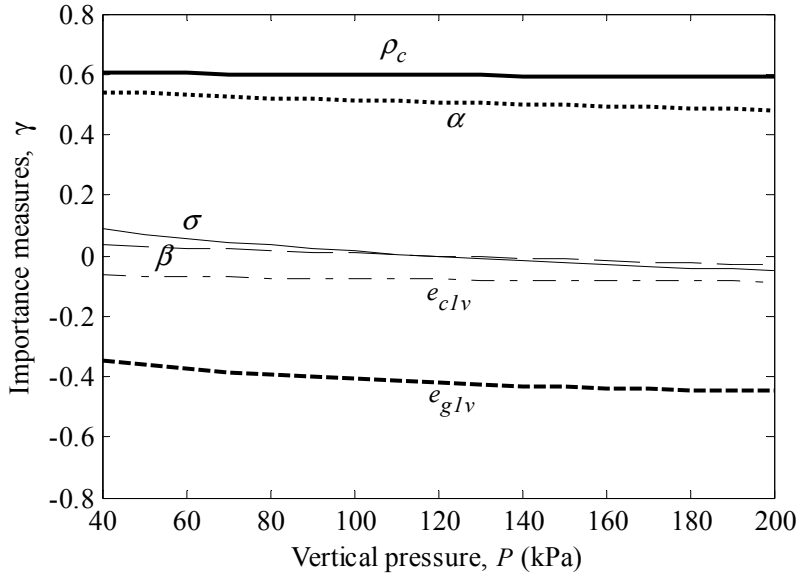
$$\boldsymbol{\gamma}^{\mathrm{T}} = \frac{\hat{\boldsymbol{\alpha}}^{\mathrm{T}} \mathbf{J}_{\mathbf{u}^*,\mathbf{z}^*} \mathbf{D}'}{\left\| \hat{\boldsymbol{\alpha}}^{\mathrm{T}} \mathbf{J}_{\mathbf{u}^*,\mathbf{z}^*} \mathbf{D}' \right\|} \tag{4.13}$$

where $\mathbf{z} = (\mathbf{x}, \boldsymbol{\Theta}) =$ the vector of the random variables, $\mathbf{J}_{\mathbf{u}^*,\mathbf{z}^*} =$ the Jacobian through which the probability is transformed from the original space $\mathbf{z}$ into the standard normal space $\mathbf{u}$, with respect to the coordinates of the most likely failure realization $\mathbf{x}^*$, $\mathbf{D}' =$ the standard deviation diagonal matrix of equivalent normal variables $\mathbf{z}'$, defined by the linearized inverse transform $\mathbf{z}' = \mathbf{z}^* + \mathbf{J}_{\mathbf{u}^*,\mathbf{z}^*}(\mathbf{u} - \mathbf{u}^*)$ at the design point $\hat{\boldsymbol{\alpha}}$. The element of $\mathbf{D}'$ are the square roots of the corresponding diagonal elements of the covariance matrix $\boldsymbol{\Sigma}' = \mathbf{J}_{\mathbf{z}^*,\mathbf{u}^*} \mathbf{J}^{\mathrm{T}}_{\mathbf{z}^*,\mathbf{u}^*}$ of the variables $\mathbf{z}'$.

Fig. 4.6 shows the importance measures of all random variables for the foundation settlement fragility over (a) S_T at a given $P = 106.5$ kPa and (b) P at a given $S_T = 450$ mm. Similar observations for the sensitivity measures can be made for the importance measures. For both the results, we can see that ρ_c is the highest important variable then α is the second highest important variable.



(a)



(b)

Fig. 4.6. Importance measures for the random variables as a function of (a) S_T at a given $P = 106.5$ kPa and (b) P at a given $S_T = 450$ mm

4.7. CONCLUSIONS

The probabilistic model for soil compression is updated to improve the accuracy of the existing model by removing a potential bias incorporating the prevailing uncertainties. The proposed probabilistic model is then used in a formulation to assess the computation of the settlement for the foundations. The settlement predictions based on the probabilistic model compare well with monitored data. Fragility estimates are developed in this study along with sensitivity and importance measures.

Fragility estimates for the settlement applied to the Treporti Test Site (TTS) in Italy show that the probability of exceedance decrease as the specified threshold settlement increase at a given the vertical pressure, and the probability of exceedance increase as the vertical pressure increases at a given the specified threshold settlement.

The sensitivity measures indicates that means of the slope of K_0-LCC regime, $\mathrm{E}(\rho_c)$, model parameter controlling the curvature, $\mathrm{E}(\alpha)$, and the reference void ratio for the K_0-LCC of the granular phase, $\mathrm{E}(e_{g1v})$, have larger effects on fragility while the means of the reference void ratio for the K_0-LCC of the clay-water phase, $\mathrm{E}(e_{c1v})$, and the model parameters corresponding to the standard deviation, $\mathrm{E}(\sigma)$ and $\mathrm{E}(\beta)$, have smaller effects over the vertical pressure increase and the specified threshold settlement.

Importance measures are also computed for the random variables in the limit state function of the fragility estimates. The similar trends can be made for the importance measures. In addition, these measures indicate that the slope of K_0-LCC regime, ρ_c, is the most important random variable for small values of the settlement

threshold then the model parameter controlling the curvature, α, becomes the second most important random variable.

CHAPTER V

RELIABILITY ANALYSIS OF TIME-DEPENDENT SETTLEMENT

5.1. INTRODUCTION

When a soil deposit is loaded, deformations of soils will occur over the time resulted in settlement. In geotechnical engineering, the total settlement can be divided into three categories including immediate, consolidation, and secondary settlement. Among these components, consolidation settlement is time dependent process in saturated fine-grained soils. During the consolidation, the changes in permeability and compressibility can affect the generation of excess pore pressure, since the rate of settlement depends on the rate of dissipation of the pore pressure. Terzaghi adopted these concepts to the consolidation theory. Then Biot (1941) extended Terzaghi's one-dimensional theory to three dimensional case and consideration of equation valid for any load variable with time. Many researchers have been proposed more general one-dimensional model based on Terzaghi's one-dimensional theory which considers the variations in compressibility, permeability and time (Davis and Raymond, 1965; Zhuang et al. 2005; and others).

However, it seems that the research on the probabilistic consolidation theory is very limited. In this paper, based on the Terzaghi's one-dimensional theory, a solution is derived for the Venice soil incorporating with finite element method.

In this study, finite element method is employed to estimate excess pore pressure. Then the probabilistic compression model developed in Chaper IV is adopted to estimate foundation settlement at different time. The fragility estimate analysis is conducted. Furthermore, sensitivity and importance measures are also computed.

The method of estimating the time-dependent settlement can be summarized as followings: (1) Compute the excess pore pressure at certain time using a one-dimensional consolidation based on a Terzaghi solution (2) Compute the settlement of the foundation at certain time using a probabilistic compression model (3) Estimate the fragility , sensitivity and importance measures using First Order Reliability Method (FORM).

5.2. NUMERICAL ANALYSIS

In general, the effective stress can be written as

$$\sigma_e = \sigma_T - u \tag{5.1}$$

where $\sigma_T =$ the current total stress, $\sigma_e =$ the current effective stress, and $u =$ the excess pore pressure.

During the consolidation, the pore pressure will dissipate due to the loading stage. Eq. (5.1) can be written as a function of time, t.

$$\sigma_e(t) = \sigma_T(t) - u(t) \tag{5.2}$$

Therefore, the time-dependent settlement depends on the portion of the excess pore pressure under time. Thus, the time-dependent settlement can be compute using the effective stress at certain time. In order to compute the excess pore pressure we employ finite element method (FEM). Then, the probabilistic soil compression model is used to estimate the time-dependent settlement.

5.2.1. Finite Element Method

The stresses due to the loading are estimated by FEM. A very simple approach has been used to represent stress-strain response of the soil. For the FEM analysis, we consider a circular loading on a linearly elastic and isotropic condition. The analysis also employs linear displacement rectangular elements and is capable of simulating an elastic modulus profile that increases linearly with depth. The steps for the FEM analysis are consist of (1) defining input data and creating mesh, (2) creating global stiffness matrix including strain-displacement matrix, constitutive matrix, and element stiffness matrix, (3) imposing of displacement boundary conditions, (4) solving for displacement with a load vector, and (5) computing the stresses.

We consider the 4-node uniform rectangular elements created from a grid of m nodes in the y-direction and n nodes in the x-direction. Each element has a horizontal dimension, a, and vertical dimension, b.

In the consolidation theory, the behavior of soil combines equilibrium equations, stress strain relations, strain-displacement relations, and continuity equation. In FEM analysis, these equations are first calculated on the element and later assembled in to global matrix. The displacement vector, $\mathbf{\Delta d}$, within an element can be formulated in terms of the displacement vector at the nodes, $\mathbf{\Delta d}_e$

$$\mathbf{\Delta d} = \mathbf{B}\mathbf{\Delta d}_e \tag{5.3}$$

where, $\mathbf{B}$ = the vector of shape function,

Then, the equilibrium equation for each element can be formulated to describe deformational behavior.

$$\Delta\boldsymbol{\varepsilon} = \mathbf{C}\Delta\mathbf{d}_e \tag{5.4}$$

where, $\mathbf{C} =$ the matrix containing the derivatives of shape function,

The vector of stress components, $\Delta\boldsymbol{\sigma}$, can be determined using the vector of constitutive , $\mathbf{D}$

$$\Delta\boldsymbol{\sigma} = \mathbf{D}\Delta\boldsymbol{\varepsilon} \tag{5.5}$$

The element matrix, $\mathbf{K}_e$, is given as

$$\mathbf{K}_e = \int_{Vol} \mathbf{B}^{\mathrm{T}}\mathbf{D}\mathbf{B}dVol \tag{5.6}$$

All elements are then combined to form the global stiffness matrix using an element stress-strain matrix.

$$\mathbf{K}_G\Delta\mathbf{d}_{eG} = \mathbf{R}_G \tag{5.7}$$

where $\mathbf{K}_G =$ the global stiffness matrix, $\Delta\mathbf{d}_{eG} =$ the vector of nodal displacements, and $\mathbf{R}_G =$ the vector of nodal forces.

The boundary conditions in the FEM simulation contain a displacement and a drainage condition. It is assumed that there are no horizontal displacements under the center of the loading. The free drainage conditions are also assumed. The stress and strain from FEM analysis can be obtained as the secondary quantities. The excess pore pressure under embankment loading was determined by Henkel (1960) for plain strain condition following as

$$\Delta u = \frac{1}{3}\left(\Delta\sigma_1 + \Delta\sigma_2 + \Delta\sigma_3\right) + \kappa\sqrt{\left(\Delta\sigma_1 - \Delta\sigma_2\right)^2 + \left(\Delta\sigma_2 - \Delta\sigma_3\right)^2 + \left(\Delta\sigma_3 - \Delta\sigma_1\right)^2} \tag{5.8}$$

where $\Delta u =$ excess pore pressure, $\kappa =$ Henkel's pore pressure parameter, and $\Delta\sigma_1, \Delta\sigma_2, \Delta\sigma_3 =$ change in major, intermediate, and minor stresses.

Terezaghi's consolidation theory is employed to describe the fluid flow through void in the soil considering the continuity of the flow. The continuity indicates that the volume of water flowing out is equal to the volume change of the water. Then, we consider the discrete problem of time domain since consolidation depends on the rate of pore pressure dissipation under the time using uncoupled analysis. After defining governing equation for an element, the assembled matrices for all elements can be obtained. The boundary conditions can be also introduced. In this study, since implicit θ method is unconditionally stable the implicit method scheme is used.

$$\left(\theta\mathbf{K} + \frac{1}{\Delta t}\mathbf{Q}\right)\dot{u}_{t+\Delta t} = \frac{1}{\Delta t}\mathbf{Q}\dot{u}_t - \theta\mathbf{K}\dot{u}_t \tag{5.9}$$

where $\mathbf{K} = \int_{Vol} C_v \mathbf{B}^{\mathrm{T}}\mathbf{B} dVol$, $\mathbf{Q} = \int_{Vol} \mathbf{H}^{\mathrm{T}}\mathbf{H} dVol$, and $C_v =$ coefficient of consolidation.

5.2.2. Excess Pore Pressure Estimates Using Finite Element Method

Since we assumed the linearly elastic, the elastic modulus profile can be written as

$$E(y) = E_0 + E_1 y \tag{5.10}$$

where, $E_0 =$ the elastic modulus at the ground surface and $E_1 =$ the rate of increase in elastic modulus with depth y.

The embankment was constructed to reach the final elevation at a 6.7 m height and 40 m diameter corresponding to about 106.5 kPa for a vertical pressure, P, at the

Table 5.1. Input variables for FEM analysis

Parameter	*value*
E_0	*35000 (kPa)*
E_0	*3000 (kPa)*
a	*0.5*
b	*0.5*
m	*150*
n	*150*
n_f	*41*
Poisson ratio	*0.15*
κ	*1*
C_v	*LN(0.2645, 0.196)*

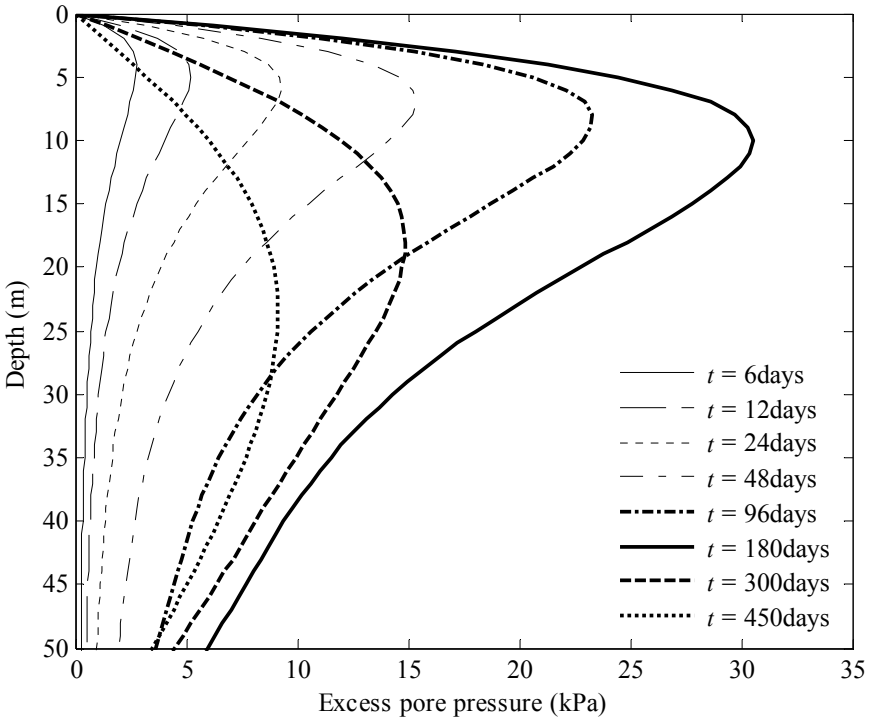


Fig. 5.1. Excess pore pressure estimates under center of the embankment

end of construction (around 180 days). The following properties shown in Table 5.1 are used in this computation. Among these parameters, the C_v has a lognormal distribution with mean and standard deviation. In order to estimate the excess pore pressure, the mean values of C_v is used.

Fig. 5.1 shows the estimated excess pore pressure under the center of loading. It can be observed that the excess pore pressure during the loading state is increased. This is caused by the loading in the new stage. However, the excess pore pressure after loading stage is dissipating continuously.

5.3. TIME-DEPENDENT SETTLEMENT ESTIMATES

In this section, we estimate the time-dependent settlement of foundation using field/laboratory data. In addition, the probabilistic compression model and the posterior statistic developed in Chapter IV for $\boldsymbol{\Theta}$ are used in this study. The detailed description of the probabilistic model and posterior statistics can be found in Chapter IV.

The estimates of the time-dependent settlement settlements are computed using the Terzaghi's one-dimensional settlement. In general, the time-dependent settlement $S(t)$ can be calculated by the summation of the individual layers, $\Delta H(t)_i$, which has varying compression properties or the initial void ratio with depth

$$S(t) = \sum_{i=1}^{r} \Delta H(t)_i = \sum_{i=1}^{r} \left\{ \frac{\Delta e(t)_i}{1+e_{0i}} H_{0i} \right\} \tag{5.11}$$

where, $\Delta e(t)_i$ = the change in void ratio under $\Delta P(t)_i - u(t)_i$ at t, $\Delta P(t)_i$ = the vertical stress increment due to a given pressure $P(t)$, e_{0i} = initial void ratio, and H_{0i} = the thickness of each soil layer.

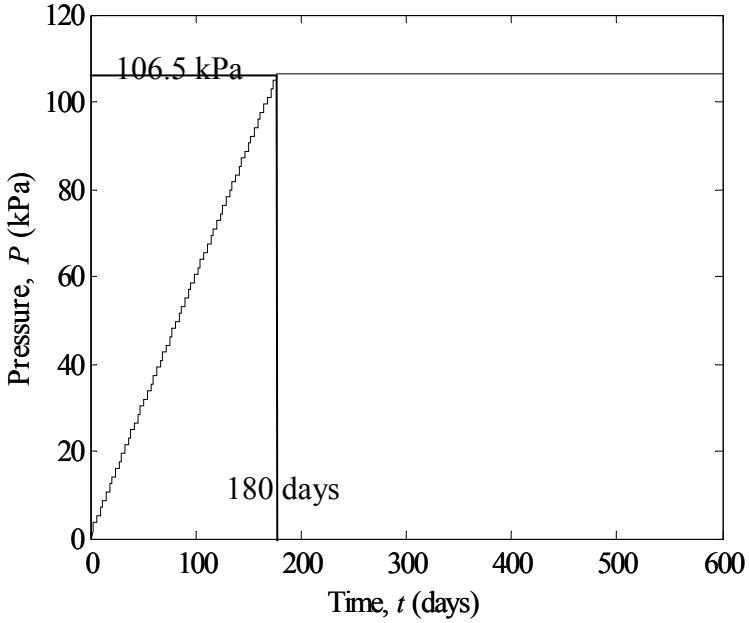


Fig. 5.2. Embankment loading with time

5.3.1. Applications to Treporti Test Site

In this section, we present an application to assess the time-dependent settlement solution. Thus, the comparison between the results based on the proposed solution and monitored data from field are shown.

As an application, we define the material properties of subsurface soil from the field observations from Simonini (2004). The load in the construction of the embankment is applied with time over a long period as shown in Fig 5.2. Simonini

(2004) provided the site investigations, the laboratory investigations, and the monitoring of the vertical displacement under the center of the embankment at different time.

In order to compute the time-dependent settlement, the model parameters $\boldsymbol{\Theta}$ and the random variables $\mathbf{x}$ are adopted from Chapter IV. In addition, we consider one more random variable C_v needed for the calculation of the time-dependent settlement. Therefore, the random variables are $\mathbf{x} = (P, e_{0i}, H_{0i}, C_v)$. As application, C_v has the lognormal distribution with mean 0.2645m^2/sec and standard deviation 0.196.

We note that each layer is assumed to be statistically independent. Therefore, the probabilistic model for the estimation of the time-dependent settlement can be written as

$$S(\mathbf{x},\boldsymbol{\Theta},t) = \sum_{i=1}^{r} \Delta H(\mathbf{x},\boldsymbol{\Theta},t)_i = \sum_{i=1}^{r} \left\{ \frac{\Delta e(\mathbf{x},\boldsymbol{\Theta},t)_i}{1+e_{0i}} H_{0i} \right\} \tag{5.12}$$

where, $\mathbf{x}_S =$ a vector of measurable variables.

$$\begin{aligned} &\mathrm{E}\{S(\mathbf{x},\boldsymbol{\Theta},t)\} \approx S(\mathbf{x},\boldsymbol{\Theta},t)\big|_{\boldsymbol{\Theta}=\hat{\boldsymbol{\Theta}}} \\ &VAR\{S(\mathbf{x},\boldsymbol{\Theta},t)\} \approx \left\{\nabla S(\mathbf{x},\boldsymbol{\Theta},t)\big|_{\boldsymbol{\Theta}=\hat{\boldsymbol{\Theta}}}\right\}^{\mathrm{T}} \boldsymbol{\Sigma}_{\boldsymbol{\Theta}\boldsymbol{\Theta}} \left\{\nabla S(\mathbf{x},\boldsymbol{\Theta},t)\big|_{\boldsymbol{\Theta}=\hat{\boldsymbol{\Theta}}}\right\} \end{aligned} \tag{5.13}$$

where $\mathrm{E}(\cdot) =$ the mean of a random variable, $\nabla S(\mathbf{x},\boldsymbol{\Theta},t)$ is the gradient of $S(\mathbf{x},\boldsymbol{\Theta},t)$ and $\hat{\boldsymbol{\Theta}} =$ the posterior means in the model parameters.

On the basis of the probabilistic compression model and Eq. (5.8), Fig. 5.3 shows the predicted time-dependent settlement using Eq. (5.12) and the monitored results in TTS. The monitored result from field is shown as solid line, the mean predicted result is shown as dash line, and the dot lines indicate $\pm\sigma_{S(t)}$. During the loading stage, predicted time-dependent settlement shows exceptionally good with monitored settlement. In this

stage, settlements are depending on both the loading of embankment and the dissipation of excess pore pressure. However, after loading (180 days), the predicted settlement is lower than the monitored settlement. In this stage, the settlement depends on dissipation of excess pore pressure.

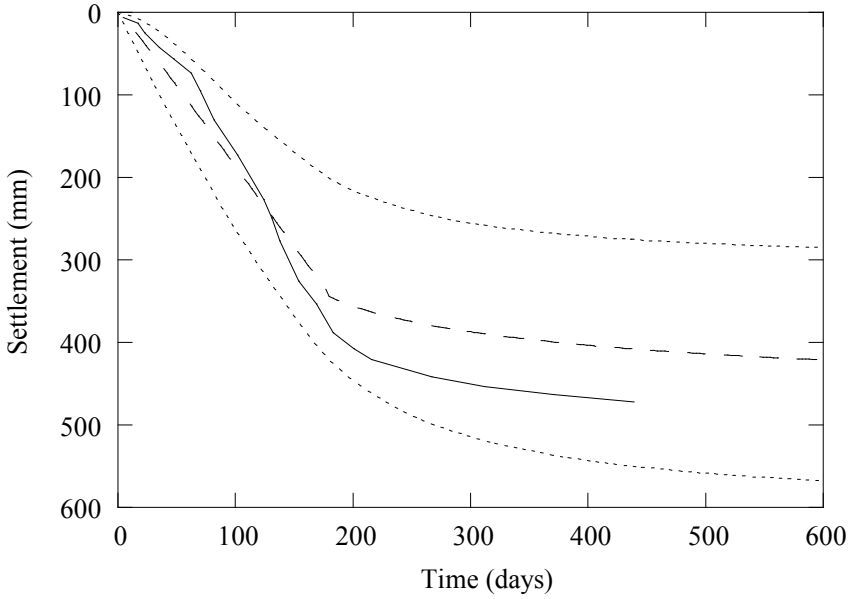


Fig. 5.3. Comparison between the predicted and the monitored time-dependent settlement

5.4. FRAGILITY ASSESSMENT OF TIME-DEPENDENT SETTLEMENT

Let $g(\mathbf{x},\boldsymbol{\Theta},t)$ be a mathematical model describing the limit state of interest. A limit state function defined such that the event $\{g(\mathbf{x},\boldsymbol{\Theta},t)\le 0\}$ denotes the attaining or exceeding a specified settlement threshold for a given value of the vertical pressure $P(t)$. Where $\mathbf{x}$ denotes a vector of measurable variables and $\boldsymbol{\Theta}$ denotes a vector of

model parameters. Using the settlement model in Eq. (5.13) and considering a specified threshold settlement, S_T, the limit state functions can be formulated as

$$g(\mathbf{x},\boldsymbol{\Theta},t) = S_T - S(\mathbf{x},\boldsymbol{\Theta},t) \tag{5.14}$$

The fragility of a foundation can be formulated as

$$F\{S_T, P(t), \boldsymbol{\Theta}, t\} = \mathrm{P}\left[\{g(\mathbf{x},\boldsymbol{\Theta},t) \le 0\} \,|\, S_T, P(t)\right] \tag{5.15}$$

where $\mathrm{P}[A \,|\, \mathbf{b}]$ denotes the conditional probability of event A for the given values of variable(s) $\mathbf{b}$. The uncertainty in the event for the given $P(t)$ arises from the inherent randomness in the variables $\mathbf{x}$, the inexact nature of the settlement model $S(\mathbf{x},\boldsymbol{\Theta},t)$ (or its sub-models), and the uncertainty inherent in the model parameters $\boldsymbol{\Theta}$.

To incorporate the epistemic uncertainties inherent in the model parameters $\boldsymbol{\Theta}$, we consider $\boldsymbol{\Theta}$ as random variables. The predictive estimate of fragility, $\tilde{F}\{S_T, P(t), t\}$, is the expected value of $F\{S_T, P(t), \boldsymbol{\Theta}, t\}$ over the posterior distribution of $\boldsymbol{\Theta}$ (Gardoni et al. 2002b), i.e.,

$$\tilde{F}\{S_T, P(t), t\} = \int F\{S_T, P(t), \boldsymbol{\Theta}, t\} f(\boldsymbol{\Theta})\, d\boldsymbol{\Theta} \tag{5.16}$$

where $f(\boldsymbol{\Theta})$= the posterior probability density function of $\boldsymbol{\Theta}$.

The fragility estimates are now a function of a specified threshold settlement, S_T, vertical pressure, $P(t)$, and the time, t. A First Order Reliability Method (FORM) is used to estimate the fragility. Fig.5.4. shows the probability of exceedance with respect to S_T at $P(t) = 106.5$ kPa. In the figure, the predictive fragilities are compared with the different time. The fragilities reflect the time-dependent settlement. An increased

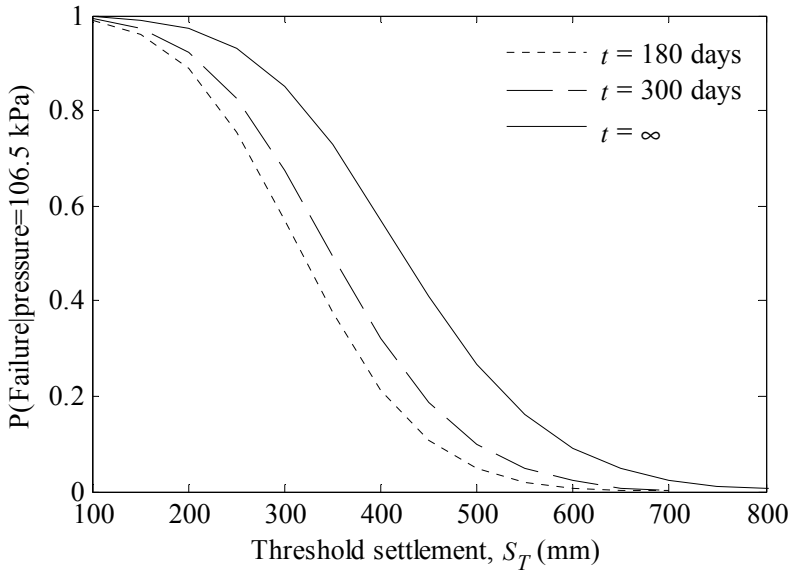


Fig. 5.4. Predictive fragility as a function of S_T at a given $P(t) = 106.5$ kPa under different times

fragility over time is can be observed.

5.5. SENSITIVITY AND IMPORTANCE MEASURES

In this section, we compute the sensitivity measures for the parameters and the importance measures for all random variable used in the estimate of the time-dependent settlement. We note that the sensitivity and importance measures are computed by using FORM.

5.5.1. Sensitivity Measures

Sensitivity analysis is employed to determine the sensitivity of the reliability of the time-dependent settlement with respect to parameters in the limit state function or in the

distribution of the random variables. The vector $(\mathbf{x}, \boldsymbol{\Theta})$ defines the set of all parameters used for the reliability analysis. We can consider the influence on β and p of unit changes in $\mathrm{E}(\boldsymbol{\Theta})$ and $\mathrm{E}(\mathbf{x})$. The sensitivity measures can be computed following Hohenbichler and Rackwitz (1986). The gradient of the first-order reliability approximation of the failure probability is obtained by using the chain rule of the differentiation as

$$\nabla_{[\mathrm{E}(\boldsymbol{\Theta}),\mathrm{E}(\mathbf{x})]} p = -\varphi(\beta) \nabla_{[\mathrm{E}(\boldsymbol{\Theta}),\mathrm{E}(\mathbf{x})]} \beta \tag{5.17}$$

where $p =$ probability of the failure estimate using FORM, $\varphi(\cdot) =$ the standard normal probability of density function, $\nabla_{[\mathrm{E}(\boldsymbol{\Theta}),\mathrm{E}(\mathbf{x})]}\beta =$ the derivative of the reliability index β with respect to $\mathrm{E}(\boldsymbol{\Theta})$ and $\mathrm{E}(\mathbf{x})$.

In this analysis, the means of the random variables are considered as the parameters for the sensitivity analysis. Since we are interested in the sensitivities of β with respect to the mean of each random variable, we need to scale $\nabla_{[\mathrm{E}(\boldsymbol{\Theta}),\mathrm{E}(\mathbf{x})]}\beta$ to compare the sensitivity measures of all parameters. On this basis we define the vector $\boldsymbol{\delta}$

$$\boldsymbol{\delta} = \mathbf{D} \nabla_{[\mathrm{E}(\boldsymbol{\Theta}),\mathrm{E}(\mathbf{x})]} \beta \tag{5.18}$$

where $\mathbf{D} =$ the diagonal matrix with diagonal elements given by the standard deviation of each random variable. We note that the vector $\boldsymbol{\delta}$ renders the element of these vectors dimensionless and makes the parameter variations proportional to the corresponding standard deviations, which are measures of the underlying uncertainties.

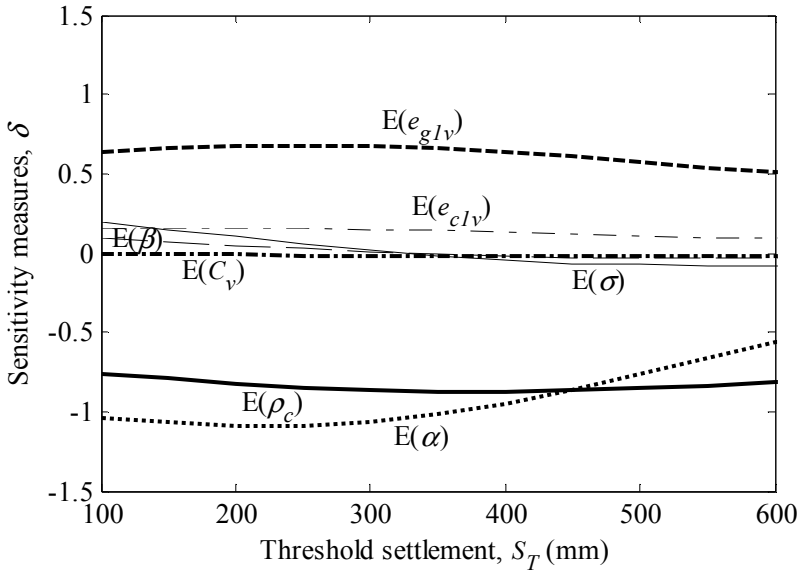
Fig. 5.5 shows the sensitivity measures as a function of S_T for $P(t) = 106.5$ kPa at (a) $t = 180$ days and (b) $t = 300$ days. We note that the sensitivity with respect to $\mathrm{E}(\boldsymbol{\varepsilon})$ is negligible and don't show them here. The results of the sensitivity measures are the same trends from the results of Chapter IV. Thus, $\mathrm{E}(\rho_c)$, $\mathrm{E}(\alpha)$, and $\mathrm{E}(e_{g1v})$ have larger effects on fragility while $\mathrm{E}(e_{c1v})$, $\mathrm{E}(\sigma)$, $\mathrm{E}(\beta)$, and $\mathrm{E}(C_v)$ have smaller effects. However, the effects of $\mathrm{E}(C_v)$ is increasing when time t increases.

5.5.2. Importance Measures

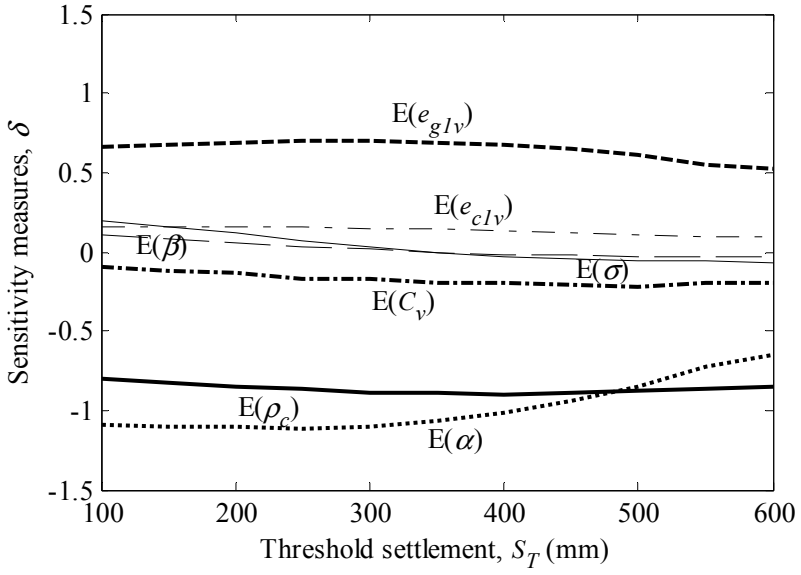
We have several random variables which have different contributions to the variability of the limit state function defied in Eq. (5.14). Following Der Kiureghian and Ke (1995), the importance measure can be defined as

$$\boldsymbol{\gamma}^{\mathbf{T}} = \frac{\hat{\boldsymbol{\alpha}}^{\mathrm{T}} \mathbf{J}_{\mathbf{u}^*,\mathbf{z}^*} \mathbf{D}'}{\left\| \hat{\boldsymbol{\alpha}}^{\mathrm{T}} \mathbf{J}_{\mathbf{u}^*,\mathbf{z}^*} \mathbf{D}' \right\|} \tag{5.19}$$

where $\mathbf{z} = (\mathbf{x}, \boldsymbol{\Theta}) =$ the vector of the random variables, $\mathbf{J}_{\mathbf{u}^*,\mathbf{z}^*} =$ the Jacobian through which the probability is transformed from the original space $\mathbf{z}$ into the standard normal space $\mathbf{u}$, with respect to the coordinates of the most likely failure realization $\mathbf{x}^*$, $\mathbf{D}' =$ the standard deviation diagonal matrix of equivalent normal variables $\mathbf{z}'$, defined by the linearized inverse transform $\mathbf{z}' = \mathbf{z}^* + \mathbf{J}_{\mathbf{u}^*,\mathbf{z}^*}(\mathbf{u} - \mathbf{u}^*)$ at the design point $\hat{\boldsymbol{\alpha}}$. The element of $\mathbf{D}'$ are the square roots of the corresponding diagonal elements of the covariance matrix $\boldsymbol{\Sigma}' = \mathbf{J}_{\mathbf{z}^*,\mathbf{u}^*} \mathbf{J}^{\mathbf{T}}{}_{\mathbf{z}^*,\mathbf{u}^*}$ of the variables $\mathbf{z}'$.

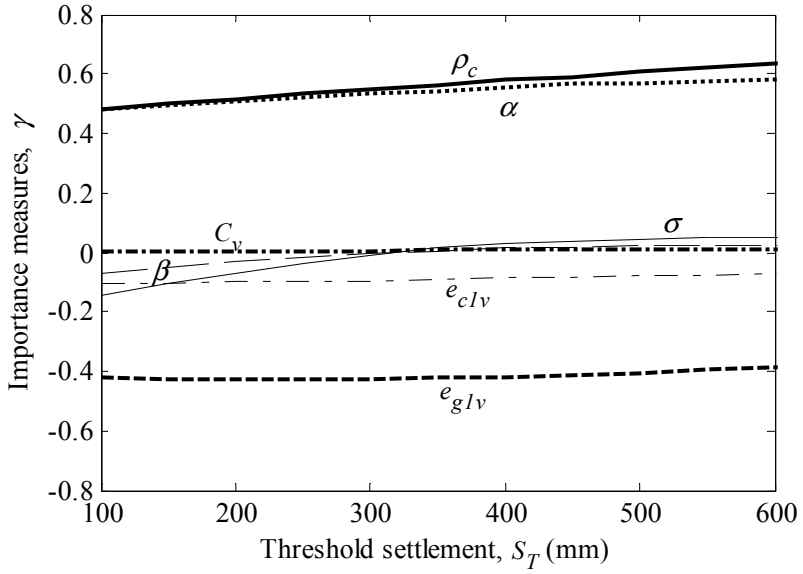


(a)

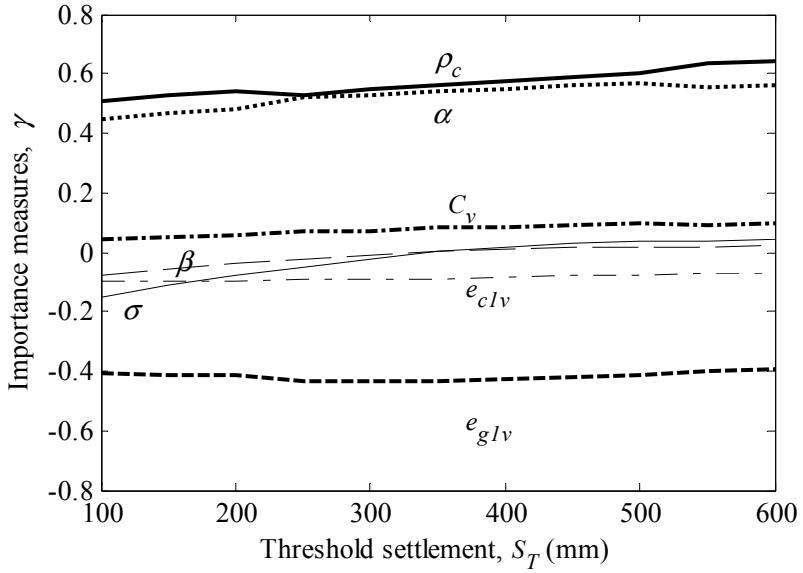


(b)

Fig. 5.5. Sensitivity measures for the parameters as a function of S_T at a given $P(t) = 106.5$ kPa (a) $t = 180$ days and (b) $t = 300$ days



(a)



(b)

Fig. 5.6. Importance measures for the random variables as a function of S_T at a given $P(t) = 106.5$ kPa (a) $t = 180$ days and (b) $t = 300$ days

Fig. 5.6 shows the importance measures of all random variables for the time-dependent settlement as a function of S_T for $P(t) = 106.5$ kPa at (a) $t = 180$ days and (b) $t = 300$ days. Similar observations defined in Chapter IV can be made for the importance measures.

5.6. CONCLUSIONS

The presented models and methodology are applied to estimate the time-dependent settlement of the foundation. Uncertainties in the probabilistic compression model and the material parameters are considered. Fragility estimates are obtained for the time-dependent settlement. It is observed that the predictive fragility keeps increasing over time.

The results of the sensitivity measures show the same trends ones shown in Chapter IV. In sensitivity measures, among the parameters, $\mathrm{E}(\rho_c)$, $\mathrm{E}(\alpha)$, and $\mathrm{E}(e_{g1v})$ have larger contributions on fragility while $\mathrm{E}(e_{c1v})$, $\mathrm{E}(\sigma)$, $\mathrm{E}(\beta)$, and $\mathrm{E}(C_v)$ have small contributions. However, the effects of $\mathrm{E}(C_v)$ is increasing when time t increases. Importance measures are also computed for the random variables in the limit state function of the fragility estimates. The similar trends can be made for the importance measures. In addition, these measures indicate that ρ_c is the most important random variable for small values of the settlement threshold then α becomes the second most important random variable.

CHAPTER VI

SUMMARY

6.1. SUMMARY

This research proposes reliability methodologies applied to Venice soil. Probabilistic soil identification is first developed with site-specific soil identification based on CPT measurements. The following conclusions are made with regard to the probabilistic soil identification:

- The proposed approach provides a simple and straightforward tool that allows updating the soil classification charts based on site-specific data;
- The site-specific classification chart can be used for a more accurate identification of the soil and accounts both for prior information available before conducting the tests and for the site-specific data;
- In practice, the developed approach can be used to construct and update regional classification charts that are be more accurate than a generic chart within the geographical area of interest;
- The information of the results can be used to produce interpretations of CPT logs, giving a profile of the probability of identifying each of the soil classes with depth.

Probabilistic soil compression model is then developed. In order to avoid dealing with different approaches for each category of soil, a simplified unified compression model is used to characterize the nonlinear compression behavior of soils of varying gradation through a single constitutive law. The Bayesian updating rule is used to

incorporate information from laboratory datasets in the computation of the statistics of the compression model parameters, as well as of the uncertainty inherent in the model. The probabilistic compression model improves the model estimates compared to the deterministic model and removes the inherent bias.

We developed a probabilistic model to estimate the settlement and the time-dependent settlement of foundations using field/laboratory data and the proposed probabilistic compression model. Fragility estimates will be obtained for a settlement by applying the proposed model. The probability of exceedance decrease as the specified threshold settlement increase at a given the vertical pressure, and the probability of exceedance increase as the vertical pressure increases at a given the specified threshold settlement. In addition, sensitivity and importance measures are computed to identify the key parameters and random variables in the model. Also, an approximate fragility estimate of the fragility is developed.

REFERENCES

Ang, A. H. S., and Tang W. H. (2006). Probability concepts in engineering: Emphasis on applications to civil and environmental engineering. 2nd edn. John Wiley & Sons. Hoboken, NJ.

Baecher, G. B., and Ingra, T. S. (1981). "Stochastic FEM in settlement predictions." *J. Geotec. Eng. Div., Am. Soc. Civ. Eng.,* 107(GT4), 449–463.

Biot, M. A. (1941). "General theory of three-dimensional consolidation." *J. Appl. Phys.,* 12, 155–164.

Biscontin, G., Cola, S., Pestana, J. M., and Simonini, P. (2007). "Unified compression model for Venice lagoon natural silts." *J. Geotec. Geoenviron. Eng.*, 133(8), 932–942.

Box, G. E. P., and Tiao, G. C. (1992). Bayesian inference in statistical analysis, Wiley, New York.

Briaud, J. L. (1997). The National Geotechnical Experiment Sites at Texas A&M University: Clay and sand: Summary, Report NGES-TAMU-007-September 1997, Civil Engineering, Texas A&M University.

Brzakala, W., and Pula, W. (1996). "A probabilistic analysis of foundation settlement." *Comput. Geotech.*, 18(4), 291–309.

Chambers, J. M. Cleveland, W. S. Kleiner, B. and Tukey, P.A. (1983). Graphical methods for data analysis, Duxbury Press, Boston, MA.

Davis, E. H., and Raymond, G. P. (1965). "A non-linear theory of consolidation." *Geotechnique,* 15(2), 161–173.

Douglas, B. J., and Olsen, R. S. (1981). "Soil classification using the electric cone penetrometer". *Proc. Cone Penetraion Testing and Experience*, ASCE, St. Louis, MO: 209–227.

Der Kiureghian, A., and Ke, J. -B. (1995). "Finite-element based reliability analysis of frame structures." *Proc., ICOSSAR '85, 4th Int. Conf. on Structural Safety and Reliability*, Kobe, Japan, 1, 395–404.

Ditlevsen, O., and Madsen, H. O. (1996). Structural reliability methods, Wiley, New York.

Elkateb, T., Chalaturnyk, R., and Robertson, P. K. (2002). "An overview of soil heterogeneity: Quantification and implications on geotechnical field problems." *Can. Geotech. J.*, 40(1), 1–15.

Fenton, G. A., and Griffiths, D. V. (2002). "Probabilistic foundation settlement on spatially random soil." *J. Geotech. Geoenviron*. Eng., 128(5), 381–390.

Fenton, G. A., and Griffiths, D. V. (2005). "Three-dimensional probabilistic foundation settlement." *J. Geotech. Geoenviron. Eng.*, 131(2), 232–239.

Gardoni, P. (2002). Probabilistic models and fragility estimates for structural components and systems. Ph.D. Dissertation, University of California, Berkeley, Berkeley, CA.

Gardoni, P., Der Kiureghian, A. and Mosalam, K. M. (2002a). Probabilistic models and fragility estimates for bridge components and systems, Report Number PEER-2002/13, Berkeley: Pacific Earthquake Engineering Research Center, University of California, Berkeley.

Gardoni, P., Der Kiureghian, A., and Mosalam, K. M. (2002b). "Probabilistic capacity models and fragility estimates for RC columns based on experimental observations." *J. Engrg. Mech.*, ASCE 128(10), 1024–1038.

Harleman, D. R. F., Bras, R. L., Rinaldo, A. and Malanotte, P. (2000). "Blocking the Tide." *Civil Enginering, ASCE,* 52-57.

Henkel, D. J. (1960). "The shear strength of saturated remoulded clays." *Proc., ASCE Research Conf. on Shear Strength of Cohesive Soils,* Boulder, CO, 533–554.

Hohenbichler, M., and Rackwitz, R. (1986). "Sensitivity and importance measures in structural reliability." *Civil Eng., Systems*, 3(4), 203–209.

Jeffreys, H. (1961). Theory of probability (3rd edition), Clarendon, Oxford, UK.

Jung, B.-C., Gardoni, P., and Biscontin, G. (2009) "Bayesian updating of a unified soil compression model." *Georisk* (in press).

Nour, A., Slimani, A., and Laouami, N. (2002). "Foundation settlement statistics via finite element analysis." *Comput. Geotech.,* 29(8),641–672.

Poulos, H. (2005) "Pile behavior–Consequences of geological and construction imperfections" *J. Geotech. and Geoenvir. Engrg.,* 131(5), 538–563.

Rao, C. R. and Toutenburg, H. (1997). Liner models, least squares and alternatives, Springer, New York.

Ricceri, G., Simonini, P., and Cola, S. (2002). "Applicability of piezocone and dilatometer to characterize the soils of the Venice Lagoon." *Geotech. Geol. Eng.,* 20(2), 89–121.

Richards, F. S. G. (1961). "A method of maximum likelihood estimation." *J. R. Stat. Soc.*, 23, 469–475.

Robertson, P. K. (1990). "Soil classification using the cone penetration test." *Can. Geotech. J.*, 27:151–158.

Robertson, P. K., Campanella, R. G., Gillespie, D., and Grieg, J. (1986). "Use of piezometer cone data. Use of in-situ tests in geotechnical engineering." *Proc. In-Situ'86, Geotechnical Special Publication No. 6*, ASCE, Blacksburg, VA:1263–1280.

Simonini, P. (2004). "Characterization of the Venice lagoon silts from in-situ tests and the performance of a test embankment." *Proc., 2nd Int. Conf. on Geotechnical and Geophysical Site Characterization*, ISC'2, 1, Porto, Protugal, 187–207.

Zeitoun, D. G., and Baker, R. (1992). "A stochastic approach for settlement predictions of shallow foundations." Géotechnique, 42(4), 617–629.

Zhang, Z., and Tumay, M. T. (1996). "Simplication of soil classification charts derived from the cone penetration test." Geotechnical Testing Journal. 19(2), 203–216.

Zhang, Z., and Tumay, M. T. (1999). "Statistical to fuzzy approach toward CPT soil classification." *J. Geotec. Geoenviron. Eng.*, 125(3), 179–186.

Zhuuang, Y. C., Xie, K. H., and Li, X. B. (2005). "Nonlinear analysis of consolidation with variable compressibility and permeability." *J. Zhejiang Univ.,* 6A(3), 181–187.

VITA

Byoung Chan, Jung was born in Korea. During 1994-1997 he served his country in the division of operation, 98th battalion, 5th Artillery brigade, the Republic of Korea Army. He enrolled at Dae-Jeon University, in Dae-Jeon, South Korea, where he received a Bachelor of Engineering degree in civil engineering in 2000. In the fall of 2002, he enrolled in the Department of Civil Engineering at Texas A&M University. He received his M.S degree in May 2005. He entered the Zachry Department of Civil Engineering program at Texas A&M University in August 2005 and received his Doctor of Philosophy degree in May 2009.

He is interested in dynamic response, constitutive modeling cohesive materials and in reliability analysis. He may be reached at his email is dodo318@gmail.com.

His permanent mailing address is:

262-2 Asan-Ri

Young-In-Myun, Asan-Si

Chung-Nam, Korea